November, 1935

50 cents per Copy
$4.00 per Year

EDUCATION

GUIDING CHILD DEVELOPMENT *By* ERNEST W. TIEGS

CONSTRUCTION OF UNITS OF WORK
By D. WELTY LEFEVER

DIAGNOSTIC TESTING IN THE UNIT OF WORK PROGRAM *By* WILLIS W. CLARK

PERSISTENCY OF ERROR AS A FACTOR IN DIAGNOSIS *By* LEO J. BRUECKNER

EVALUATION IN THE SECONDARY SCHOOL
By J. MURRAY LEE

RECENT DEVELOPMENTS IN EDUCATIONAL MEASUREMENT *By* DAVID SEGEL

SOME PROBLEMS IN MEASUREMENT
By ERNEST O. MELBY

OLD PEDAGOGUES IN NEW PASTURES
By L. F. ROVER

EDUCATION

HERBERT BLAIR, Editor-in-Chief
Professor of School Administration
Boston University School of Education

ASSOCIATE EDITORS

ARITHMETIC
GUY M. WILSON
Professor of Education
Boston University

ART
ROYAL BAILEY FARNUM
Director, R. I. School of Design

CLASSICAL LANGUAGES
B. L. ULLMAN
Professor of Latin
Chicago University

COMMERCIAL
FREDERICK G. NICHOLS
Professor of Commercial Education
Harvard Univ. School of Education

CREATIVE WRITING
TRENTWELL MASON WHITE
Lecturer in Creative Writing
Curry School of Expression

ELEMENTARY EDUCATION
AMY BRONSKY
Assistant Professor of Elementary Education, Ohio State University

ENGLISH
EVERETT GETCHELL
Professor of Education
Boston University

EXTRA CURRICULAR ACTIVITIES
EDWARD J. EATON
Professor of Education
Boston University

GEOGRAPHY
DOUGLAS C. RIDGLEY
Professor of Geography in Education
Clark University

GUIDANCE
JESSE B. DAVIS
Dean, School of Education
Boston University

HISTORY
DANIEL C. KNOWLTON
Professor of Education
New York University

HOME ECONOMICS
GLADYS B. JONES
Director, Garland School of Home Making, Boston

HEALTH EDUCATION
MABEL C. BRAGG
Associate Professor of Education
Boston University

INDUSTRIAL ARTS
ARTHUR B. MAYS
Professor of Industrial Education
University of Illinois

MEASUREMENT
ERNEST W. TIEGS
Dean of University College
University of Southern California

MUSIC EDUCATION
JAMES L. MURSELL
Professor of Education
Teachers College, Columbia University

PHYSICAL EDUCATION
FREDERICK RAND ROGERS
Dean of Physical Education and Health
Boston University

SCIENCE
HANOR A. WEBB
Chairman, Div. of Science and Math.
George Peabody College for Teachers

SECONDARY EDUCATION
FREDERICK E. BOLTON
Professor of Education
University of Washington

SOCIAL SCIENCE
ERNEST R. GROVES
Professor of Sociology
University of North Carolina

SUPERVISION
ERNEST O. MELBY
Dean, School of Education
Northwestern University

VISUAL EDUCATION
F. DEAN McCLUSKY
Director, The Scarborough School

Address all communications to the office of the publishers

THE PALMER COMPANY, Publishers

120 Boylston Street Boston, Massachusetts

NEWCOMB & GAUSS CO., PRINTERS, SALEM, MASS.

EDUCATION

A MONTHLY MAGAZINE

DEVOTED TO THE SCIENCE, ART, PHILOSOPHY AND LITERATURE OF EDUCATION

HERBERT BLAIR, EDITOR-IN-CHIEF
PROFESSOR OF SCHOOL ADMINISTRATION, BOSTON UNIVERSITY

VOL. 56 NOVEMBER, 1935 No. 3

MEASUREMENT NUMBER

ERNEST W. TIEGS, EDITOR
DEAN OF UNIVERSITY COLLEGE, UNIVERSITY OF SOUTHERN CALIFORNIA

CONTENTS

		PAGE
Guiding Child Development Through Diagnosis and Measurement.	*Ernest W. Tiegs*	129
Pertinent Problems in Teaching a Unit of Work	*Lillian A. Lamoreaux*	130
The Place of Measurement in the Construction of Units of Work.	*D. Welty Lefever*	132
Diagnostic Testing in the Unit-of-Work Program	*Willis W. Clark*	138
Persistency of Error as a Factor in Diagnosis	*Leo J. Brueckner*	140
A Contribution of Measurement to Remedial Reading	*Blanche L. Garrison*	144
Evaluation in the Modern Secondary School Classroom	*J. Murray Lee*	148
The Use of Scholastic Aptitude Test Data in Elementary School Educational Administration	*Alfred S. Lewerenz*	154
Recent Developments in Educational Measurement	*David Segel*	161
The Construction and Standardization of the Knauber Art Tests.	*Alma Jordan Knauber*	165
Some Problems in Measurement	*Ernest O. Melby*	170
Limitations of Freedom	*Ridgley C. Clark*	175
Summary of Published Research Studies in the Field of Latin Teaching.	*Lillian Dale Thomas*	179
Old Pedagogues in New Pastures	*L. F. Rover*	186
Department of Book Reviews		190
New Educational Books		192

Indexed in the Readers' Guide and in the Education Index

Published by THE PALMER COMPANY
120 Boylston Street Boston, Massachusetts

RAYMOND P. PALMER, Business Manager

Published monthly except July and August. Subscription price $4.00 per year. 50 cents per copy. Entered as second class matter at the Post Office, Boston, Mass., under the Act of March 3, 1879.

Tyrrell's American History Exercises

By JAMES F. TYRRELL, Master of the Thomas J. Kenny School, Boston

True-False Exercises *Completion Exercises*

Recognition Exercises

For daily lesson assignments, home work and examinations.

These exercises are in line with the modern thought of leading educational authorities as to the need of supplementing the old familiar type of examination with something more comprehensive and yet less laborious and time-taking.

Section I—The Period of Exploration.
Section II—The Settlement of Virginia.
Section III—The Settlement of Plymouth.
Section IV—The Settlement of Massachusetts Bay.
Section V—The Settlement of New York.
Section VI—The Settlement of New Hampshire and Connecticut.
Section VII—The Settlement of Maryland and Rhode Island.
Section VIII—The Settlement of Delaware and The Carolinas.
Section IX—The Settlement of New Jersey and Pennsylvania.
Section X—The Settlement of Georgia.
Section XI—Struggle between France and England for the possession of North America.
Section XII—1763-1783.
Section XIII—1783-1829.
Section XIV—1829-1865.
Section XV—1865-1898.
Section XVI—1898-1935.

Each of the above sections contains separate sheets of each type of examination. The three types of exercises covering one section should be given a few days apart, thus allowing the teacher to emphasize between times the points that are giving trouble.

These exercises will interest the pupils. They are more comprehensive and can be given in less time than the old type of examination. They can be marked accurately and uniformly. Teachers will be able to check their instruction and discover the weak points.

TYRRELL'S AMERICAN HISTORY EXERCISES are sold in packages of 150 sheets on each of the above sections. A package contains 50 sheets of each of the three types of exercises, or enough for fifty pupils.

Price, per package (specify which section).....$1.35
Smaller quantities at 1 cent per sheet. Minimum 25 cents
Complete sample set .....60 cents

The Palmer Co. 120 Boylston St., Boston, Mass.

EDUCATION

Vol. 56 NOVEMBER, 1935 No. 3

GUIDING CHILD DEVELOPMENT THROUGH DIAGNOSIS AND MEASUREMENT

ERNEST W. TIEGS
UNIVERSITY OF SOUTHERN CALIFORNIA

PLANNED first as an entire number on the Evaluation of Outcomes of Units and Activities, it was soon evident that while we talk and write much about progressive education, much of our educational practice must still be described in more traditional concepts. However, a number of offerings are distinctly related to recent progressive trends, while others are related to other equally important aspects of evaluation.

No discussion of diagnosis or measurement can be complete unless it deals with other aspects of the problem of guiding development, including aims or objectives, materials, and procedures. Accordingly we have asked Lillian Lamoreaux to set the stage in her article on "Pertinent Problems in Teaching a Unit of Work." The contribution of Dr. Lefever clarifies many problems met regularly by those attempting to develop valid and effective units and activities. Mr. Clark explains specifically how to use carefully developed diagnostic tests in making units yield the largest returns in child development. Dr. Brueckner's study on the persistence of error throws a strong and not too favorable light on the reliability of most of the diagnostic work done up to the present time. Blanche L. Garrison sounds a new note on the secondary level in her contribution on remedial reading in the senior high school. Dr. Lee's helpful article on evaluation of outcomes on the secondary level will do much to unite those who are enthusiastic about progressive procedures, and those who also insist on progressive outcomes. The title of Mr. Lewerenz' article suggests that it would be of interest primarily to administrators, but it should be read by every teacher as a piece of significant research in guidance.

Everyone is interested in what David Segel says about measurement. His broad grasp of the many factors involved and his opportunity for viewing many aspects of the field serve to give his readers perspective. The problem of test validation is particularly difficult in the field of fine arts. Workers in measurement will be interested in how Alma Jordan Knauber attacked the difficulties involved.

Dean Melby of Northwestern University presents a fine summary of the present situation in measurement. It should be helpful to both radicals and conservatives in the movement.

PERTINENT PROBLEMS IN TEACHING A UNIT OF WORK

LILLIAN A. LAMOREAUX
SUPERVISOR OF ELEMENTARY EDUCATION
BURBANK CITY SCHOOLS, BURBANK, CALIFORNIA

MUCH has been done in the last few years to aid the teacher in building units of work; teachers have been taught to build the unit and have been aided in the acquisition and selection of materials. Less has been said and done to help the teacher in the techniques in presenting units of work. Many teachers who are successful in building an excellent unit fail entirely when it comes to carrying it through with a class. This paper is an attempt to present a few of the difficulties met by teachers in putting into effect units of work. These difficulties are grouped under four headings, namely: (1) Initiating the unit; (2) Handling groups; (3) Balancing types of activities; (4) Evaluation.

Teachers are taught, when building a unit, that they must provide for many approaches in order that they may have more than one means of arousing and stimulating interest and creating an earnest desire to know more of the subject at hand. The environmental influences affecting learning could well receive more attention. If it is true that all learning takes place through the interaction of the learner and his environment,[1] then the teacher should be more alert to the type of environment she provides, not only in initiating the unit, but all through the carrying forward of the unit to its completion.[2] The teacher should be cognizant also of the fact that she, herself, is part of the environment.

It is not uncommon for a teacher to say to her principal or supervisor: "I have made a beautiful approach to our work; the children are interested, but just how do we get into the work? What will I do with the things the children have brought? How can I get my pupils into the real study of our problems?" The trouble with this teacher is that she does not know how to plan with the children. Teachers need definite help, not only in long-time planning with the children but in short-time and immediate planning. Long-time planning, as here used, is conceived of as the proposed steps to be used in the solution of the problem; it makes up the total solution pattern. Short-time planning may be the week to week plans which carry forward, adjust, or modify the long-time plans. Immediate plans take care of single, definite steps toward the solution of the problem. They help the children to know definitely what is to be done at the present period. Example: A trip to the harbor is scheduled. The children planned to take with them their crayons and sketching boards, which consisted of stiff cardboards with manila paper clipped to them. Each child was to choose a boat and make a sketch of it.[3]

1 Thorndike, E. L.: "Human Learning."
2 Porter, Martha Peck: "The Teacher in the New School." Page 27.
3 Los Angeles City Schools—Course of Study Department. Bulletin 299. Page 3. Boats, 5-6.

The handling of groups presents many vital problems—problems which may make or mar the success of the work. These problems often determine the happiness of the teacher and pupil in their work, the discipline in the room, and, to a great extent, the learning resulting from the work. It is here the teacher asks such questions as:

"On what basis shall I form groups or committees? Do all committees work under the direction of a chairman? Who should be the chairman? What are his functions? Should the children work together on one project in the group, or should there be individual projects in the group? How can I keep several groups busy at one time? Where will I get materials for the groups? Should I provide it or should the children bring it?"

"I have four committees in my room. We make good plans, but there is such confusion when they start to work. How can I overcome this?" The problem here is one of releasing committees. It is the confusion at this point that causes persons to misjudge the entire program and also the teacher. A teacher needs to be helped to a better organization and placing of materials in order to dispel confusion. A method of releasing one group at a time may tide a teacher over until the pupils gain a sense of self-responsibility.

Sometimes teachers fail to see their responsibility after the planning is done and the committees are formed. Teachers are responsible for guidance all through the work. They should meet with committees; solicit advice from the whole class when a committee needs it; contribute and aid in the selection of books, pictures, and other materials; raise questions which guide thinking; encourage those who need it; aid in the adjustment of individuals to their group and their work; etc. The alert teacher is a guidance director at all times.

Keeping the class informed as to what the groups are doing and how they are progressing offers many problems.

"How often shall we make reports to the class? Who should do the reporting? Should all the reports be written and filed in a notebook? What will I do with a child who only looks at pictures and will not read? What will I do with John? He fools away his time and does not have much evidence of work to show or report, while Don works at home because he is so interested. He does all the work for the group. Should I hold my groups together, or should I let some finish ahead of others? If some finish first, what will I do with them until the rest catch up?"

Along with the matter of reporting comes the problem of holding the entire class responsible for reports made on work done in other groups. A class working on a problem in Communication had groups working on early attempts in communication, wireless, radio, telephone, etc. The teacher was much perturbed over the fact that they were not all doing some reading on all the topics. This was a real problem to her, and she needed definite help in securing some general learning on the total problem.

Everyone working with the unit of work today is faced with the problem of balancing types of activities. So keenly is this problem felt that many

are working out some type of child accounting card on which is shown the units experienced and the types of activities participated in in each unit. Many are realizing that it is possible for a child to go from unit to unit experiencing the same type of activity, especially if the child is left to choose his own. Should provision for the balancing of types of activities be made from unit to unit, or is it not possible, by giving it more thought, to balance in some degree the activity experiences of the child within each unit?

Some of the most interesting and most worthwhile problems arise in connection with the thought of evaluation. We are quite in agreement that we do want an evaluation of attitudes, skills, and knowledge. A new idea has entered educational thought as to who should evaluate. In the past it was the teacher's province and problem; today it is the child's. We now recognize him as the doer, the learner, and the one who should be most vitally interested in his own progress.

When should we evaluate? How shall we evaluate? What is the relation of evaluation to planning? These are other vital problems. Planning and evaluation are both fundamental parts of experience. Since evaluation affects future planning, much thought and attention should be given to making it as meaningful as possible.

THE PLACE OF MEASUREMENT IN THE CONSTRUCTION OF UNITS OF WORK

DR. D. WELTY LEFEVER
ASSOCIATE PROFESSOR OF EDUCATION
UNIVERSITY OF SOUTHERN CALIFORNIA

MANY educators, enthusiastic over progressive trends in curriculum organization and teaching methods, have under-estimated the place and importance of measurement in modern education. This attitude has resulted in part from a previously misplaced emphasis on the mechanical features of testing and in part from the fact that tests are not as yet valid for evaluating many of the more subtle goals of education. Measurement, however, if carefully and intelligently employed, has much to offer those who are striving for a new reality in the educative progress.

Without a definite basis for determining the nature and extent of the child's growth, educational planning will be reduced to sheer guesswork. Even though we freely admit the many criticisms of educational measurement, there still remains the necessity for setting specific as well as general goals, immediate as well as remote objectives. A systematic study of each child with reference to these outcomes is the only alternative to the dissipation of educational effort.

The process of planning units of pupil activity demands a thorough knowledge of the instruments and techniques of educational measurement. At best the information now obtainable concerning the nature of a pupil group constitutes an all too meagre basis for organizing a curriculum. We know so little about capaci-

ties, specific talents and defects, drives and incentives, pupil interests and personal objectives. In spite of the fact that we stand in serious need of additional facts concerning the personnel of our pupil population, the sources of knowledge now available are far too significant to be ignored.

Obviously these units of work are learning situations and it is of the greatest importance that they be organized in the light of this fact. The individual pupils' capacity to learn is a critical factor in the situation as a whole. All learning is of necessity individual learning and is conditioned by that individual's ability to learn. Our group and individual intelligence tests through experiment and observation have shown themselves to be quite valid indicators of the power to learn. This is especially true when the learning has had to do with verbal materials. This phase of mental capacity is much more significant than it may seem at first thought. We often think of verbal responses as superficial and relatively unimportant. However, the underlying ability to deal successfully with abstract and verbal symbols is basic to all scholarship and intelligent human progress.

Mental age constitutes a valuable index of the child's experiential maturity. Naturally the intelligence test is not a complete measure of the child's experience since the items composing the test are selected because every child taking the test could be reasonably expected to have become familiar with them. The degree of intellectual maturity indicated by the child's mental age is closely correlated with the level of complexity of the tasks which the child can be expected to perform successfully. In constructing a unit of work the average complexity of the activities involved should be carefully considered in the light of the average mental age of the group.

Of still greater importance than the group average is the group range. This principle obviously applies to all measures employed to guide the planning of a unit of work. Among the activities possible within the unit of work must be some simple enough for the lowest mental age represented. Other activities should challenge the most mature capacity in the group. Research is greatly needed to determine the nature of tasks suited to the several mental age levels.

The intelligence quotient is considered by most students of individual differences to be a fairly valid measure of the rate of learning of which the child is capable. It is probably not ar valuable a measure for curriculum planning as mental age except for those children who represent extremes in chronological age. If a dull older child and a bright younger child have the same mental age, the type and complexity of the activities anticipated from each should be distinctly different. The bright child will probably be interested in discovering explanations and causes; the dull child will be limited to the more obvious aspects of the unit. Abundant opportunity should be provided in the plan for all types of intelligence.

Much light may be cast on the planning of units of pupil activity and experience through a study of the learning processes involved. The success or failure of these activities depends entirely on the nature and extent of

the pupil learning which results therefrom. Significant progress has been made in the study of learning curves for different type situations. The construction of a curriculum should be influenced by a consideration of such factors as positive or negative acceleration in the learning curve and plateaus in the curve due to the reorganization of bonds, loss of interest, or to the increased complexity of the problem. Use may also be made of the results found by research in the economy of learning.

The work of Courtis relative to growth curves may well suggest valuable ways of studying learning in various fields of activity. In determining the progress of the learning process measurement is of necessity the major tool.

Studies of maturation with animals and with children indicate that certain types of learning will take place more efficiently at particular levels of maturity. More definite measures of this factor are sorely needed, although chronological age and mental age are approximate indices of maturation. Valuable research has been completed yielding information concerning physical growth and maturity. Social maturity is a more subtle factor and little success has so far been attained in measuring it. Perhaps no field of study will furnish more important assistance to the curriculum maker than the development of measures of maturation whether physical, mental, or social.

In order to correlate the materials to be employed in building a unit of work with the abilities, skills, experience, background, and interests of the pupil group, it is necessary to measure a number of characteristics of those materials. Questions relating to the difficulty of materials can be answered best by an experimental approach involving a careful count of errors and of the frequency of avoidance by members of the group. Again a measure of average difficulty is not sufficient. It is more important to know the range. The complexity of analysis required of each child in the group will, of course, vary with the type of participation in the activity expected of him. Scales for indicating at least the approximate degree of intricacy of the mental tasks involved are needed. Such scales might well be related to the maturation and index of brightness of many children.

Outstanding among the various types of difficulty to be considered in assembling materials for a unit of work is the problem of reading skill and background required of each pupil. Reading is obviously a tool of critical importance in most pupil activities. It consists of many elements, one of which relates to the vocabulary employed in the various items of material requiring pupil reading. Much valuable work has been accomplished in measuring the occurence of the different words used. Difficulties inherent in the ideas or concepts back of many simple words have not been adequately gauged as yet. Easy words may represent complex ideas requiring experience on the adult level for complete understanding. Vocabulary diversity can be measured and should be considered as indicating the possibilities for pupil progress in acquiring new vocabulary. The availability of materials suitable for developing the vocabulary of all the pupils par-

ticipating in the study is important since the range of vocabulary grade placement in most elementary or high school groups has been shown to be great.

A relatively definite test of the level of reading comprehension can be made for the group. Of course, such a measure could apply only approximately to the many types of reading possible in a single unit of work. We frequently take too much for granted when we assume, for example, that a child who reads poetry and narrative easily will handle science material with equal facility. If possible, a diagnostic survey of reading difficulties should be made so that incidental opportunities for remedial reading may be provided. Freeing the child from a constant awareness of the mechanics of reading does not relieve the teacher of the responsibility for providing abundant opportunity for indirect remedial drill.

A careful coordination of reading materials with the reading grade placement of the group is a prime necessity in curriculum organization. In selecting text books, reference books, supplementary reading, recreational reading, incidental clippings, magazines and the like, the range of reading comprehension should be continually in mind.

Since many units of pupil activity presuppose a large amount of independent study or so-called "research" on the part of the child, it is naturally important to know the reading difficulty of the reference materials. Much of this type of material, such as encyclopedias, dictionaries, maps, tables of facts, source-books, biographical dictionaries, and other similar reference books have never been analyzed for reading difficulty. Of equal practical value would be a measure of the "get-at-ableness" of this material for different types of pupils. How complex are the tasks we expect children of varying abilities and experience to perform as incidental to our unit of work?

Some units of work involve numerical concepts, arithmetical computation, and arithmetical reasoning. The ability to do the necessary computation and the difficulty of such computation as is required is relatively easy to measure. More difficult to evaluate is the reasoning ability of the pupil as demanded by various problem situations. Still more subtle to measure and control are the number ideas or concepts as they appear in the pupil's activity. It is probable that these number concepts are of greater fundamental value for the pupils understanding of the world in which he lives than all of the computation so heavily stressed. In constructing many units of work these number concepts will have a natural and important place. They should be studied by the curriculum maker.

All pupil activities assume certain backgrounds of information and pupil experience. The intelligence test is an approximate measure of the general extent of this experience. Much more definite and specific knowledge is needed in order to determine the pupil's readiness to begin particular activities. It is possible to make fairly definite and diagnostic surveys of pupil information in the several fields of study. How much, in general, as measured by a careful sampling, does the child know about his-

torical events and personages, social or civic organizations, geographical facts or scientific facts and principles? What specific knowledge is required for a child to participate successfully in a given unit of work? Naturally, this participation may take place on different levels of ability and background knowledge.

Many incidental skills are involved and are assumed in the organization of a unit of work. In order to do as may be expected of him, each pupil must possess certain minimum levels of skill-efficiency. These levels of skill will vary in harmony with the quality of achievement anticipated from each pupil as gauged by his capacities and interests. Fairly satisfactory tests are available for determining the level of efficiency for many skills. For example, we may measure spelling ability, composition quality, the legibility of handwriting, the correctness of the language usage employed by the child and other incidental tools of learning.

Many important outcomes and factors relating to a unit of work are not measurable by any means now at hand. It seems to be true that the more significant a factor is the more subtle and difficult it is to measure. Where subjective judgment is required, that judgment may have varying degrees of worth. Subjective judgment may be improved by making careful, painstaking observations as a basis for judgment. These judgments should be as specific as possible. A grave danger implicit in the exaggerated enthusiasm of some "progressive educators" lies in the use of many vague generalizations in describing the work of the pupil and its outcomes. These specific judgments will not be superficial or meaningless if they are unified by a basic philosophy of education.

Subjective judgment can be improved by making it more critical. We must endeavor to include negative as well as positive observations. It is easy to see only the successes and to overlook the short-comings in an experimental program. To find the faults and weaknesses in a unit of work has greater constructional value than locating its achievements.

Subjective judgment may be improved by observing in a regular and systematic manner the factors and outcomes too evasive for direct measurement. If the profession can come to any reasonable degree of agreement concerning the meaning of the words or terms employed to describe the phenomenon under observation, the judgments of a careful observer should have some significance.

A few of the more subtle objectives and elements to be considered in constructing a unit of work will be enumerated here. All pupil activities will involve appreciation values. These outcomes should receive definite planning in the organization of the unit. Opportunities for pupil experiences and judgments concerning art and music constitute essential elements in many units. No unit of pupil activity lacks the possibility of stimulating some type of appreciation in the pupil. He may learn to value the sacrifices of our pioneer forefathers, to realize how much skill and patience the inventors required, or to thrill at the marvels of the world of science. He may begin to experience the fascination of discovering for himself some

of the basic principles controlling the life processes and the phenomena of nature.

Systematic and specific judgments should be made by the curriculum builder relative to the socializing values of each pupil activity. So little is known about this type of outcome and so much is glibly claimed by some educators that a careful study of this problem is clearly indicated. What specific types of social responses do different pupils make to the learning situations presented in the unit of work? Comparatively little research has yet been done in the field learning in a social matrix.

Much has been said and little is known about pupil interests. How do they arise? To what extent should the activity seek to arouse and guide pupil interest? In how far should the activity conform to the pattern of pupil interests already present?

Experiment and research will doubtless make available to the curriculum builder through the slow processes of growth more successful measuring devices for analyzing certain phases of pupil experience at present too complex for direct measurement. We need tests that are more analytical, that indicate specific attainments and difficulties. Measures of reasoning power and the ability to understand fundamental relationships are slowly evolving. Studies of social maturity and its interdependence on mental growth and chronological age will shed valuable light on certain problems of curriculum construction and teaching technique. For example, should pupil groups be formed on the basis of chronological age?

While these much-needed measures are being developed, the curriculum builder can and should employ the tests and measuring devices now available. Only through their careful and systematic use can he bring into mutual harmony the pupil group and opportunities for valuable pupil activity.

DIAGNOSTIC TESTING IN THE UNIT-OF-WORK PROGRAM

WILLIS W. CLARK
DIRECTOR OF ADMINISTRATIVE RESEARCH
LOS ANGELES COUNTY SCHOOLS

THE informal, activity, or unit-of-work instructional program is receiving wide approval among progressive educators. Originating in special classes and private schools, this type of educational program and its concomitant philosophy, is affecting the curriculum and the teaching procedures throughout the grades. With proper safeguards, the writer believes that this program represents a distinct advance in educational practice.

Under the formal program (characterized by the use of logical content in basic text-books, prepared syllabi, and frequent tests), there has been the assurance that basic skills and facts would be presented to all pupils. In an informal situation, with teacher and pupil freedom in selection of materials for study, there is some danger of the omission of essential elements of basic skills and subject-matter.

The solution of this difficulty lies in the frequent and intelligent use of diagnostic testing techniques. This will enable the teacher, as director and guide, to understand the difficulties of each pupil and to assure herself that the pupils are attaining reasonable mastery of techniques and orientation in subject-matter suited to their capacity and maturity.

The statement has been made, "Now that we have activity programs, we won't need any more tests." The opposite is true. If teachers are to assist in the direction of educational experiences so that individual differences and developmental needs of pupils will be cared for, a much more extensive testing program will be required. In fact, the use of diagnostic tests should be considered one of the professional techniques of the teacher.

The tests for such use may be informal and unstandardized, or they may have been prepared by those who have made an analysis of the learning elements in the given situation. The teacher should have the ability to administer and interpret tests of capacity, maturity, and accomplishment in her particular area.

There are certain assumptions carried over from the formal graded educational regime which hinder effective teaching and learning activities. One of these relates to the meaning of the word "grade". To many persons, it implies "the place where pupils learn specific things." It should be defined as "an administrative device for classifying pupils for the convenience of instruction." Rather than assume that a pupil who has completed a certain grade has a mastery of certain skills and basic information, we should admit and take into account the fact that usually less than one-half of the pupils in a given grade have the ability and accomplishment set up for that grade. Many pupils have less

and many have more than the median achievement which is normally set up as the standard for the grade.

Likewise, many school districts give but lip service to the demonstrated facts relating to individual difference in capacity. Backward pupils are being provided with an enriched curriculum when the instructional materials are brought down to the level of their comprehension and experience. A similar enrichment is provided for intellectually superior children when they are freed from needless drill in skills already mastered, and allowed to develop individual initiative and assume responsibility suited to their developmental needs.

These situations may be illustrated by data secured by use of the Progressive Achievement Tests.[1] These tests are diagnostic tests in the fundamental skills of reading, arithmetic, and language which are essential for progress in the academic aspects of education. These tests are so constructed to provide the following types of information:

1. The customary age or grade placement results for each subject.
2. A profile indicating the strengths and weaknesses of each pupil in the elements of the skills tested.
3. A diagnostic analysis of the teaching-learning situations whereby the teacher may group pupils, give individual help, and otherwise direct an intelligent educational program suited to the needs of her pupils.

The following outline, giving the elements of Arithmetic Reasoning and Addition of Arithmetic Fundamentals of the above-mentioned test, Elementary Battery, will illustrate the practical values of such information for a teacher who is attempting to individualize her instruction:

Arithmetic Reasoning.

A. Number Concept:
—Writing numbers.
—Writing money.
—Roman numbers.
—Concept of whole numbers.
—Concept of fractions and decimals.

B. Signs and Symbols:
—Signs.
—Abbreviations.

C. Problems:
—One-step.
—Two-step.
—Sharing and averaging.
—Square measure and cubic content.
—Percentage.
—Ratio.

Arithmetic Fundamentals.

D. Addition:
—Simple combinations.
—Carrying.
—Zeros.
—Column addition.
—Adding money.
—Denominate numbers.
—Adding numerators.
—Reducing fractions to common denominators.
—Adding mixed numbers.
—Adding fractions and decimals.
—Writing decimals in columns.

Some of the implications for teachers are shown by test results as follows:

1. "Two thousand six" or "ten thousand eight" are written correctly by 4th grade, 79%; 6th grade, 89%; 12th grade, 79%.

2. Knowledge that $\frac{2}{3}$ is largest of series $\frac{2}{3}$, $\frac{1}{4}$, 1/5, 2/5, and that $\frac{3}{4}$ is largest of series of $\frac{2}{3}$, $\frac{5}{8}$, $\frac{3}{4}$, 4/9 for various grades are as follows: 4th grade, 13%; 6th grade, 31%; 12th grade, 77%.

3. Knowledge that 50% is less than 2 is shown by 69% of 12th grade pupils. 31% of 6th grade pupils have the same information.

1 Tiegs, E. W., and Clark, W. W.: "Progressive Achievement Tests, Primary, Elementary, Intermediate, and Advanced Batteries." Los Angeles: So. Calif. School Book Depository.

It is obviously true that many of the problems of adjustment and success in the various grades is due to the fact that pupils do not have the mastery of certain skills or facts which they are assumed to have. On the other hand, much time and energy is expended in repeating that which is already known.

By the extensive use of well-selected measuring devices and the intelligent interpretation of the results, we may expect the unit-of-work program to provide a means for a superior type of educative activity. Without such aids, the activities are apt to fail to provide some of the essential elements for learning and thereby result in the programs not attaining its potentialities in fostering improved educational outcomes.

PERSISTENCY OF ERROR AS A FACTOR IN DIAGNOSIS

LEO J. BRUECKNER
UNIVERSITY OF MINNESOTA

VARIABILITY of response is an important factor in educational diagnosis that is often not taken into account in analyzing the results of arithmetic diagnostic tests. It is commonly assumed that a pupil's performance in a particular test given under certain conditions is a reliable index of his performance in other trials of the same test given at some subsequent time. The reliability of tests is usually determined by comparing the relative positions of the pupils' scores in the distributions of scores. This procedure fails to take into account the fact that while the two scores of an individual in the test may be practically the same, the specific items incorrect on the two tests and the kinds of errors made may be quite different.

An analysis of the specific errors on one performance of a test may, therefore, lead to a wholly invalid diagnosis. The results of a series of experiments bearing on this point are herewith reported.

The first experiment deals with the question of persistency of error when pupils repeat the identical example a number of times. Six difficult types of examples in division of fractions were mimeographed on a slip of paper. The 24 pupils in a sixth grade class that had studied the topic were directed to work these six examples and then to hand in their work. As each pupil turned in his paper, he was at once given another slip containing the same six examples and was told to work them again and then to turn in his paper. This procedure was repeated until each pupil had turned in six slips. No effort was made to motivate the work, nor was the purpose of the experiment explained to them. The pupils were not told at any time whether or not their work was correct.

An analysis of the six papers of each pupil revealed the extent to which incorrect solutions of each example persisted in the six trials. In all there were 24 times 6 or 144 oppor-

tunities for making from 0 to 6 errors on six repetitions of the same example. In 96 of these 144 opportunities it was found that there were no errors on any of the six repetitions of the same example. This means that in 48 cases there were errors on one or more of the six trials of the same example. Table I shows the detailed analysis of these results.

TABLE 1

Analysis of Incorrect Solutions in Six Trials of Each Example.

No. of Incorrect Solutions in Six Trials	Frequency	Per Cent
1	19	39.6
2	5	10.4
3	6	12.5
4	2	4.2
5	3	6.2
6	13	27.1

In 19 cases there was only one incorrect solution in the six trials of an example; in 5 cases there were errors in two of the six trials; in 6 cases there were errors in three of the six trials; in 2 cases there were errors in four of the six trials; in 3 cases there were errors in five of the six trials; in 13 cases the particular example was worked incorrectly in all six trials. It is thus clear that in many cases there is great variability in the performance of pupils on repeated trials of the same examples. In only 33.3 per cent of the cases in which there were errors was the same example solved incorrectly in five or six of the six trials. In 50 per cent of the cases there were errors in only one or two of the six trials of the same example. These errors distributed themselves over all six trials. Similar results were found when the experiment was repeated with other groups using time intervals of an hour, a day, and a week between the test periods. It is thus apparent that the pupils' performance on a single trial of an example is not a sure index of his performance on subsequent trials of the same example.

Another aspect of this problem deals with the extent to which a pupil's performance in one example of a given type is an index of his ability to solve other examples of the same type. By "type" we mean an example involving a particular combination of skills. It is easy to construct a number of examples illustrate this idea of type. Both involve the same skills but different numbers:

$2/3 \times 1\frac{1}{2} = 1 \qquad 2/5 \times 2\frac{1}{2} = 1$

A test exercise containing four examples of each of six different types in multiplication of fractions was constructed. The 24 examples were arranged in a random order on a single sheet. The pupils worked all of the examples in one class period, as a regular class exercise. They were given as much time as they needed to complete the work. In all 327 pupils in five different cities were tested. The median number of examples solved incorrectly was 5.4; the range was from all correct to all incorrect. In 1101 cases, from 1 to 4 of the four examples of a given type were worked incorrectly. Table 2 shows the distribution of errors.

TABLE 2

Distribution of Incorrect Solutions of Six Types of Examples (4 Examples of Each Type).

Number of Incorrect Solutions	Frequency	Per Cent
1	443	40.2
2	261	23.7
3	159	14.4
4	238	21.6

In 40.2 per cent of the cases in which there were errors in one or more of the four examples of a given type, the pupil solved only one of the four examples incorrectly; in 23.7 per cent of the cases there were two incorrect solutions; in 14.4 per cent of the cases there were three incorrect solutions; in 21.6 per cent of the cases all four examples of the particular type were solved incorrectly. These results show that a pupil may work incorrectly an example on a test, consisting of a series of single examples, each of a different type, and yet work correctly other examples of the same type.

In the 238 cases in which there were incorrect solutions on all four examples of a given type, an analysis of the kinds of errors in each example yielded significant information regarding the persistence of the kinds of errors that were made by the pupils. In only 16.0 per cent of the cases were all four examples of a given type solved incorrectly for the same reason, that is, the same error was made in all four examples. In 27.8 per cent of the cases there were found, in addition to a constant error in all four examples, other errors in from one to three of the four examples. In 26.5 per cent of the cases three of the examples were missed for the same reason, often accompanied by some other error, but a different error was found in the fourth example. In 22.7 per cent of the cases two of the four examples were missed for the same reason, but each of the other two examples was missed for a different reason. In 7.0 per cent of the cases each of the four examples was solved incorrectly for a different reason.

These facts show the likelihood of incorrect diagnosis when the teacher bases her conclusions on an analysis of the pupil's performance on a single example of a given type. Elsewhere the writer has given specifications for constructing diagnostic tests that will lead to valid diagnosis.[1]

It is interesting to speculate on the causes of the variability of error revealed by the above experiments. It is obvious that various conditions may affect the pupil's performance. For example, the score of a third grade pupil on a rate test of knowledge of addition combinations may be low because of slow speed of writing rather than because of lack of knowledge of number combinations. When the same pupil is given as much time as he needs, he may be able to write all of the sums correctly. In some cases when speed is stressed there is a breakdown of skill due to lack of control of the various mechanisms involved, which under other conditions function satisfactorily. Some pupils make higher scores on tests when speed is stressed than they make on the same tests when they are given as much time as they need to complete the test. Rewards of various kinds, such as gold stars or money, praise, and awareness of growth affect the performances of pupils in different ways. It is, therefore, very difficult to get a satisfactory measure of a pupil's ability by a single test given under a certain set of conditions, since under other conditions the results may be quite different. Ratings of pupils whose scores are equal on a test given under one

1 Brueckner, L. J., and Hawkinson, M.: "The Optimum Order of Arrangement of Items in a Diagnostic Test." Elementary School Journal, Vol. 34 (January, 1934), 351-358.

set of conditions often differ considerably when the test is given under slightly different conditions.

An illustration of the ways in which the effort put forth by the pupil may affect his test scores is given by Courtis.[2] Table 3 lists the scores of two fourth grade pupils on six trials of a test of addition combinations given under different conditions, except that the time allowed on all trials was the same. The different conditions are given as the lower part of the table.

TABLE 3

Scores on Six Trials of a Test of Addition Combinations Given Under Varying Conditions.

Trial	*1*	*2*	*3*	*4*	*5*	*6*
Pupil A.						
No. attempted	30	37	30	32	37	42
No. correct	30	27	29	32	35	33
Pupil B.						
No. attempted	50	53	40	30	52	54
No. correct	49	52	39	29	51	54

Instructions to pupils:

Trial 1. See how many you can add correctly in the time allowed.

Trial 2. Work as fast as you can. Disregard accuracy.

Trial 3. Be sure every answer is right. Disregard speed.

Trial 4. See if you can work as you did on the first trial.

Two days later:

Trial 5. See if you can make as high scores as you did the other day. (Praise.)

Trial 6. I will give you a nickel if you make a higher score. (Money reward.)

The striking thing about the performances of these two pupils is the variation in their scores on the six tests. On trial 1 Pupil A attempted 30 combinations and had all of them correct. On trial 2 when speed was emphasized he attempted 37 combinations but had only 27 of them correct. Apparently the greater effort put forth broke down controls conditioning accuracy. On trials 3, 4, and 5 there was considerable variation in the number attempted but the work was quite accurate. On trial 6 when the money reward was offered Pupil A attempted more items than on any of the other trials, but his work was highly inaccurate, as it was on trial 2. Some factor that breaks down the skill was operating. It should be clear that on all trials the basic arithmetic ability of the pupil was constant. His test performances on the basis of which to evaluate his ability were affected by the ways in which he responded to the various conditions under which the test was given. The greater effort made on trials 2 and 6 may have broken down the basic controls.

The results for Pupil B do not reveal any tendency for his ability to break down under any of the test conditions, since on all trials his work was very accurate. However, there was a wide variation in the number of items attempted on the six trials of the test. His score on trial 2 was considerably higher than on trial 3 as was to be expected; his score on trial 4 was much lower than on any of the other trials, possibly due to loss of interest on the fourth trial of the same test. On the next day his score on the test under the stimulus of a promised reward was higher than on any of the other trials. In the case of Pupil B the greater amount of effort put forth on trials 2 and 6 did not break down the controls in-

[2] Data from an unpublished study.

volved as happened with Pupil A. Courtis reached the conclusion that scores on "single trials of educational tests cannot be reliably interpreted without some knowledge of the effort level of response."

Similar variability in pupil performance has been revealed by our researches in spelling and English composition. Our findings suggest that diagnosis is not as simple and mechanical a task as it is sometimes assumed to be.

Research on the various ways in which such factors as differences in motive, degree of effort, level of maturity, and nature of controls are related to pupil achievement and performance on tests, offers a field of investigation that should yield rich returns to those interested in problems in learning.

A CONTRIBUTION OF MEASUREMENT TO REMEDIAL READING

BLANCHE L. GARRISON
THOMAS JEFFERSON HIGH SCHOOL, LOS ANGELES

FIVE years ago senior high school teachers thought that their students, having completed the ninth grade, had mastered the tools of learning sufficiently well to do regular high school work in the traditional academic subjects. The last few years have been disillusioning because the results of objective tests have proved that quite the opposite is the case.

This was found to be true at the Thomas Jefferson High School in Los Angeles. It was found in the A10 English classes, there were boys who were repeating the work for the third and fourth semesters. The course of study at that time included considerable formal grammar, and the students had to pass a series of tests before they could earn A10 English credit. They couldn't read well enough to understand the text. Not more than twelve in any class were up to grade in reading, and their grade placements ranged from third to tenth, with about two thirds of them falling in the fifth and sixth grade groups. It was not strange that they could not do regular high school English.

An experimental class was established the next term. Sixteen of the worst cases were given an extra reading class in addition to their regular schedule. No home work was required, but remedial reading procedures were followed. The results were gratifying and seemed to indicate that remedial instruction would be profitable for such students.

A remedial reading program was worked out on a more scientific basis, and during the past four years it has more than justified the conclusions drawn from the hasty and entirely unreliable experiment with the sixteen students.

A counsellor visits each junior high school in the district to give the Sangren-Woody Reading Test to the A9 pupils. The students entering Jefferson who do not achieve a placement of seventh grade or better are programmed for a remedial reading class.

Those who are sixth grade in reading achievement are placed in one group; those below sixth are placed in another. Seventh grade ability is set as the standard because most of the newspapers and magazines of a popular type are of that degree of difficulty.

Within the first week the purposes and objectives of the class are discussed. The whole matter of remedial reading is considered. All students are given the results of their reading tests part by part, and they plot their own reading profiles on the chart provided for them on page two of the tests. These profile charts are of great benefit in motivation, and they are valuable to the teacher for diagnostic purposes.

The September to February, 1935, class was a representative group. Table I presents the data for this group.

With two exceptions these 33 people were all sixth grade in reading grade placement. There were 15 negroes, 7 Mexicans, 6 Italians, 1 Chinese, and 4 Americans. The racial and language difficulties in the group were obvious. The I. Q.'s ranged from 65 to 107 with a median of 85. There were only 4 with I. Q.'s of 95 or above, while 14 were below 85. This was another obvious reason for their reading troubles.

Part of the reading program is to increase the number of words read per minute. Table II indicates the reading rate for this class at the beginning of the term at the end of ten weeks, and at the end of the twentieth week.

The early gains in speed continued through the term, but a speed of 250 words appears to be the limit for the average remedial reading pupil. The scores on speed tests for several groups sustain this fact.

During the first few weeks of the semester the teacher strives to gain the confidence of these pupils and interests them in some kind of reading. The approach is based upon the students' interests, and aims toward extensive rather than intensive reading.

TABLE I

Reading Scores and Speed Scores at the Beginning and at the End of a Semester for a Typical Remedial Class.

Pupil	Race or Nationality	I. Q.	Sangren-Woody Test, Form A	Sangren-Woody Test, Form B	First Speed Test	Second Speed Test	Third Speed Test
C. J.	Italian	80	6.5	5.7	118	...	222
D. B.	Negro	88	6.9	7.5	...	162	168
D. A.	American	71	6.3	6.3	143	...	165
E. F.	American	82	6.6	6.6	194	186	162
E. J.	Mexican	73	6.6	7.4	250	344	258
G. J.	Mexican	93	6.0	8.5	200	230	293
J. T.	Negro	107	6.6	8.5	239	210	241
L. F.	American	65	6.0	8.5	214	204	241
L. D.	Chinese	91	5.4	7.0	171	256	293
L. F.	American	85	6.0	8.5	220	254	265
L. P.	Italian	70	6.2	7.3	162	185	231
L. F.	Mexican	82	6.7	7.7	118	154	195
M. V.	Negro	84	6.4	7.0	115	150	199
M. E.	Negro	72	6.0	7.1	117	138	175
P. J.	Italian	84	6.1	8.2	154	208	206
R. M.	Negro	85	6.2	8.5	195	180	225
R. E.	Negro	80	6.0	8.5	117	140	175
R. L.	Negro	71	4.9	7.2	260	250	278
R. W.	Negro	92	6.9	8.5	173	192	230
R. C.	Mexican	85	6.0	7.0	164	180	212
S. J.	Negro	90	6.5	8.5	201	159	218
S. M.	Mexican	83	6.6	8.2	146	184	221
S. E.	Italian	95	6.7	8.5	145	186	234
T. C.	Negro	83	6.3	8.5	187	212	222
V. F.	Mexican	88	6.0	8.5	188	205	284
V. P.	Mexican	85	6.0	8.5	230	240	226
W. L.	Negro	92	6.3	8.1	117	129	154
W. M.	Negro	80	6.1	7.0	120	184	282
W. C.	Negro	85	6.5	8.3	...	215	282
W. J.	Negro	101	6.6	8.5	236	279	320
T. L.	Negro	84	6.2	7.4	...	183	...
N. T.	Italian	93	6.0	6.2	170	...	...
F. F.	Italian	99	6.5	8.5	235	271	284

TABLE II. *Reading Rates.*

	Low Score	High Score	Median
Test 1	115	260	171
Test 2	129	244	184
Test 3	154	320	231

They bring any books they may choose in which to do free reading. Each person reads a paragraph or two aloud to the teacher. This serves a threefold purpose: it gives her an opportunity to observe the type of books read; it gives her a chance to observe reading habits, such as pointing out words and bad eye movements; and it provides a natural situation in which the pupil and teacher may become better acquainted. By following this procedure, the attitudes of these slow and very limited groups is greatly improved. This approach, coupled with the motivating influence of the profile charts, practically eliminates discipline problems.

The regular weekly program of the class consists of three days of reading from a common book, one day of oral reading of poetry or plays, and one day of free reading. Once a month on the free reading day one of the public librarians comes to review books.

Mondays are utilized for the oral reading of poetry and plays. Both types of literature are meant to be read aloud, so that an artificial situation is not created. The class begins with poetry. One verse is written on the board, and the class decides what the thought groups are. Then each student selects a poem of his own, studies it, and finally reads it for the group's enjoyment. The class decides, in each case, whether the thought of the poem has been made clear by the manner in which it was read. Individual help is given, not so much by the teacher as by the other students. Sometimes the reader re-states the thought in his own words; sometimes the rhythmic qualities are noted especially. These people of limited ability do enjoy "entertaining" each other once they lose the fear that comes from too frequent failure. From the first they are given only work in which they can succeed.

This semester the two classes have been especially eager to work with plays. The difficulty has been to get suitable one-act plays in simple enough language. Here, as in poetry, reading for thought and proper word grouping is stressed.

On Tuesdays, Wednesdays, and Thursdays a common textbook is used. The class members read silently about thirty pages of easy material each day. During the following class period an objective type test is given—usually true-false. The sole reason for the test is to make certain that the students are reading at least thirty pages a day. This amount sometimes crowds them, but the crowding forces them to read more rapidly. The day before an assignment is made the possible word difficulties are discussed and clarified. These are limited to not more than ten. On these days, too, a very brief mimeographed exercise giving directions, or stating a simple arithmetic problem, or relating a very obvious puzzle is passed to each student, and he writes the answer on his test paper.

The class first reads *Jim Davis*, then *Man for the Ages, Wild Animals I Have Known, Up from Slavery, Ancient Man,* and others. While reading *Up from Slavery,* they select the central thought of each paragraph; while in *Ancient Man* the order of events is emphasized, thus giving drill in organization.

Fridays are reserved for a good

time. Library books are read and enjoyed for recreational purposes. The public librarians in the district have cooperated in making an annotated book list of appropriate books. The books on this list are placed on special shelves in each of the branch libraries, and the remedial reading people are urged to make themselves known to the librarians so that they may be given special consideration. These librarians have been reviewing books over a period of three years, and they feel that their efforts have been worthwhile. They report that some of the pupils in the first remedial reading groups have continued to use the library. This continued influence in leisure-time reading is, of course, the most valuable feature of this work.

In addition to this cooperation from the public librarians, the school librarian assists very materially. She has used the fine money to buy books appropriate for the remedial reading classes. There are now available some 50 to 75 easy and interesting books. These are on a special shelf so that the students may more easily select what they desire. Ten recreational books a semester is the goal set for these people, but they average only five. A very simple book report is filled out and a block graph showing each pupil's record is posted in the room.

Do these people improve in reading ability under this regime? Measured objectively, they do. At the end of the semester Form B of the Sangren-Woody Test is given. In the same class of 33, 3 people failed to make any improvement; 1 scored less on the second test than on the first; 16 made seventh or eighth grade placements; and 13 were above eighth grade placement, thus making an improvement of at least two years in their reading grade in twenty weeks' time. Translated into percentages, these figures indicate that about 39% can hope to do average work in such subjects as English and social studies, that 48% can do very mediocre, but passing work, and that 3% will probably fail. These students working at their best can only keep up with the average pupil. If some personality clash exists between them and the teachers, they are almost certain to lose out. If the environment for such pupils is not conducive to study, there is further loss. It is only under the most favorable circumstances that they can attain the average accomplishment in academic subjects.

Remedial reading work offers opportunities in several fields of endeavor. In the tests and measurements, more scientific procedures for classroom use are yet to be made practical. Control-group technique in high-school classes is not often used. A study of the correlation between reading improvement and improvement in general school achievement for these particular groups should be made.

For the person interested in writing there is much to be done in providing wholesome reading material of suitable psychological maturity.

The stereopticon can be used to flash jokes, properly phrased current event material, and guessing games on the screen. Some experimenting in this last field has been done, and such procedures will be profitable when enough material has been accumulated.

Finally, the remedial reading work

relieves the teacher of many of her problems. Because the pupils are interested and because there is an almost unbelievable improvement in the classwork, the teacher knows that her efforts are worthwhile. It is much easier to maintain an enthusiasm for these retarded groups when the pupils and teachers both realize they are doing something worthwhile.

EVALUATION IN THE MODERN SECONDARY SCHOOL CLASSROOM

J. Murray Lee
DIRECTOR OF RESEARCH
BURBANK CITY SCHOOLS, BURBANK, CALIFORNIA

Place of Evaluation

EVALUATION has a definite place in the teaching process of the modern secondary school classroom. If one searches the literature dealing with progressive teaching procedures in the secondary schools, one is apt to feel that evaluation is playing the role of the "forgotten man." This is partly due to the fact that one group is trying so hard to convince teachers that there are better ways of teaching that they have had little time to consider the evaluation of what they are doing; and the other group, the measurement experts, have not realized that a different approach is required. Actually, it is possible for the groups to agree upon a plan of evaluation which is satisfactory to both. The main purpose of this article is to "bridge the gap" for the teacher who is trying to utilize the modern methods and is a little puzzled about the place of evaluation.

The planning of the unit which is to be taught should include a listing of the aims of education which the teacher hopes to develop further through the unit. The learning experiences are then set up so as to realize the aims. Evaluation enters into the process to determine to what extent the aims have been realized. One great difficulty with many teachers is that they do not formulate clearly the aims which they are trying to attain. In such cases those units are apt to "wander far afield," and accomplish little. No wonder such teachers feel that evaluation is out of place. They have nothing to evaluate.

Once the teacher has formulated the aims of the unit, half of the job of evaluation is completed. For the evaluation is to be made in terms of the aims. This concept of measurement is markedly different from the concept which results in making so many pages of the textbook the basis of a test. Many object, and rightly so, to the testing for isolated unrelated facts, but no one can object to careful attempts to measure the achievement in terms of the aims of the course.

Illustrations will probably make this principle of evaluation much clearer. Such material will give also the opportunity to show what techniques of measurement are needed and how they operate. The basis used for this development is the *Tentative Course of Study for the Core Curriculum of Virginia Secondary*

Schools — Grade VIII.[1] In the Virginia course of study the aims of education have been presented in somewhat detailed lists of "Emotionalized Attitudes, Generalizations, and Special Abilities." From these lists the teacher selects the aims for each unit of work. It is essential that the teacher, in carrying on the work of the unit and the evaluation process, consider the aims in terms of the level to which such aims should be developed for the grade taught.

Method of Evaluation

In our present stage of development, evaluation can best be made by utilizing two tools of measurement, tests and judgments. Tests can be constructed which will measure many of the attitudes, generalizations, and abilities, though not all of them by any means. The others, which cannot be measured by tests, must be measured by judgments. There are two types of activities with which judgments may be concerned. One is the judgment of a pupil's reaction or performance which would be an indication of the extent to which an attitude was being changed, a generalization developing, or an ability increasing. The other would be the judgment of a product or an exhibit or a completed piece of work which would show growth. The only difference between the two is that criteria must have been developed for evaluating the reaction before it takes place, while with the second such criteria can be developed after the product is completed.

The three initial steps are the same for the construction of tests or the development of criteria for utilizing judgment. It is only in the fourth and last step that there is a difference between the two. The four steps in the development of evaluative instruments are as follows:

1. Formulate the aims or objectives of the course or unit. In the Virginia course of study, this step requires the selection of the attitudes, generalizations, and abilities which the unit of work is expected to develop. A number of these aims will be given later when samples of tests are presented.

2. Divide the aim into workable parts. For instance, the "Ability to Read" can be divided into these parts:

   The ability to use the mechanics needed in reading.
   The ability to analyze, interpret, and evaluate reading materials.
   The ability to use reference books.
   The ability to interpret thought in oral reading.
   The ability to use silent reading as a recreation.

It can be clearly seen that the subdivisions under the main ability are much more helpful as guides for evaluation than the statement of the "Ability to Read" taken alone.

3. Decide what behavior, reaction, or response is expected to show whether the aim has been attained. This step requires the interpretation of the aim in terms of pupil responses. That this step is a most valuable training device for teachers is at once apparent. Most teachers seldom consider the types of learning situations which they expect will result in the attainment of the aims. This type of approach in forcing teachers to consider their units in relation to the aims is concrete and requires more than mere "lip service" on the part of

1 Published by the Virginia State Board of Education, Richmond, Virginia, 1934, 319 pp.

the teachers. An illustration of the procedure is in interpreting "the ability to use reference books" in terms of responses. The student, to have acquired the ability to use reference books, would need to be able to:

Locate books in the library.
Use general reference works.
Use a table of contents.
Use an index.
Read the material located to get answers to problems.

4. It is in this fourth step that the procedure differs. For *tests* it is necessary to construct test items to show to what extent the behavior is present. A sample table of contents could be given and questions developed which would show whether the pupil could use the table to locate material.

For *judgments* it is necessary to develop criteria for evaluating the performance, product, or reaction to indicate to what extent the desired behavior is present. These will consist of a number of points which should be utilized in making the particular type of judgment necessary. Such statements tend to refine the judgments and make them more reliable.

Evaluation by Means of Tests

A more complete understanding of the process of evaluation can be attained by studying examples of such materials. In this section, there are a number of partial tests which were developed to measure definite aims and in the next section examples are given of means of utilizing judgments. The tests in this section were all developed by teachers to measure various aims in their units.[2] It should be understood that the tests are merely small samples of much longer tests and only enough items are given to illustrate the possibilities. Tests, rather than judgments, should be used when it is possible to use them, and the clever teacher will find it is possible to measure many more aims than was once thought possible. For each example, the aim, the directions, and a few sample items are given.

Unit on Science for eighth grade by George Wood.

Aim: Ability to avoid laboratory accidents.

Place a check by the best answer.

30. The safest way to break off a piece of glass tubing of a desired length is:

....1. Break it as you would a stick.
....2. Strike it with a metal instrument.
....3. Mark it with a file and break opposite mark.
....4. Wrap it in a cloth.
....5. Hold it under water.

32. In order to avoid an explosion resulting from lighting a hydrogen generator:

....a. Be sure the air is out of the generator.
....b. Wrap a cloth around the generator.
....c. Tie the lighter on the end of a long stick.
....d. Put the whole apparatus under a hood.
....e. Stand at a distance from the generator.

Unit on Science by T. Epps Jones.

Aim: Ability to collect, care for, and preserve plants.

Rate the best job—5, rate the next best—4, the next best—3, the next best or next to the worst—2, and the worst job—1.

B. Some flowers, such as "Pansies," are very fragile; they bruise easily, lose color and shape readily. Several pupils in the class collected them.

2 Constructed by Virginia teachers in the author's class in "Measurement of Learning" in the College of William and Mary during the summer of 1935.

a. Some put them in weak solution of alcohol. ....
b. Some preserved them in plain water. ....
c. Some pressed them between blotters. ....
d. Some pressed them between sheets of waxed paper. ....
e. Some pressed them between sheets of plain paper. ....

D. In preparing to transport some small bushes which were already taken from the ground, the students did the following things:

a. Some shook all of the soil from the roots, so that the plants would be lighter and easier to carry. ....
b. Some carried the plants just as they were carried from the ground. ....
c. Moistened the roots and wrapped each plant in an old piece of bag. ....
d. Some shook part of the soil from the plant and placed it in a box. ....
e. Some moistened the roots of the plants and put several in a bucket. ....

Unit on Social Science for eighth grade by W. D. Barr.

Aim: Ability to discriminate between honest and misleading advertising.

Mark the statement *I* if it represents an impossible claim on the part of the advertiser; *F* if the statement represents a claim so unreasonable that it is likely to be false; *T* if the statement represents a claim that seems reasonable and is likely to be true.

....1. *X* will restore your youth..
....2. Millions prefer *X* cigarettes.
....3. *X* toothpaste ends forever pink toothbrush.
....4. *X* crystals are the only way to health.
....5. *X* will reduce your weight six pounds overnight.

Beside the measurement of various abilities, it is also possible to construct tests which will measure application of principles and generalizations. The following samples are taken from a test in the process of standardization and are reproduced with the permission of Miss Gladys Omohundro.

Application of Principles.

Check the correct answer and the reason for the answer which you give:

Two pieces of ice of equal size are selected. One is wrapped in white cloth and the other in black cloth. They are placed side by side in the sun. Which will melt more quickly? Why?

Answers:

1. The piece of ice wrapped in the black cloth will melt first.
2. The piece of ice wrapped in the white cloth will melt first.
3. Both will melt in the same length of time.

Reasons:

1. White cloth will not receive or take in heat like the black cloth.
2. The white cloth will reflect the heat while the black cloth will absorb the heat.
3. The white cloth will not hold the cold or heat.
4. The black cloth absorbs light.
5. The black cloth reflects heat and the white cloth absorbs light.
6. The black cloth is tighter than the white cloth and does not let air in on the ice as much.
7. The white cloth attracts heat from the sun and the blach cloth does not.
8. Black absorbs more heat than any other color.

Generalizations.

Check the best generalization:

A piece of iron pipe weighing five pounds was corked at each end and put on the water; it floated. Another similar piece of iron pipe weighing the same was put on the water without being corked. It sank.

1. A piece of pipe displacing more water than its own weight will float.
2. The cork in each end of the pipe caused more water to be displaced and thus caused the cork to float.
3. The pipe with the cork in the end floated because it had a large air cell in it which made it buoyant.

4. In the first iron pipe, the amount of air per square inch inside of the pipe was greater than the amount of water per square inch, so the pipe floated. In the second case this was not so.

5. Air is lighter than water.

6. The amount of water displaced by the second pipe was equal to its own weight.

7. The density of iron is greater than the density of water, so it sinks.

8. The water has not enough buoyancy to hold the iron up.

Attitudes can also be measured by using similar methods. The following samples were constructed to measure certain social and economic attitudes. The first is on the eleventh grade level and the second on the eighth.

Check the answer which most nearly expresses your opinion. (Constructed by Ivan Butterworth.)

II. The Townsend Plan for Aged People.

1. The government should not subsidize old age and people should make provisions for their own welfare in later years.
2. The government should aid old people in some form of insurance where aid is needed.
3. The government should furnish all people with this plan of insurance, regardless of individual need.

XV. Huey Long and his plan of Share the Wealth.

1. His policies are not in keeping with American ideals.
2. Where enormous wealth has been in the hands of the few, it should be gradually taken by the government.
3. All wealth should be confiscated and distributed.

Answer Yes or No to the following questions. (Constructed by Ella Goldstein.)

....9. Do you believe that foreigners should be allowed to send money to their native homes for their families who are still there?

...15. Do you believe that immigrants should be segregated in certain sections until they have become Americanized?

Evaluation by Means of Judgment

Judgments should supplement tests in the evaluative process. There are many factors for which tests have not been built and it is in evaluating these factors that the role of judgment is useful. It is necessary to develop lists of criteria as means of judging the reactions of the pupils to various learning experiences. There are two types of results which need to be considered. One is a product or exhibit which has been constructed and the other is a reaction or performance of the pupil.

Teachers have long been utilizing judgments, but in most cases many factors are omitted for the teacher does not think of them at the moment. Using lists of criteria has the advantage of including all of the factors necessary in the forming of a complete judgment. This point is well illustrated in the lists which follow.

An example of criteria for judging a product in the following list used in evaluating a frieze which a group had drawn to illustrate the evolution of transportation. The questions were:

1. Does it show a knowledge of the development of transportation?

2. Does it show an understanding of man's increasing control of nature?

3. Does it represent proper interpretation of reading materials?

4. Is it pleasing composition?

5. Does it show an increasing use of the ability to use color?

6. Does it represent an increasing expression of individuality?

A careful study of these questions show that the frieze was evaluated

from the viewpoint of understandings developed (questions 1 and 2), abilities increased (questions 3, 4, and 5), and attitudes evolving (question 6).

An example of a basis for judging observed reactions is found in this list of criteria for evaluating oral reports. The first six questions deal with the method of presentation and the other three with the organization of the material.

1. Is there free flow of thought?
2. Does he use clear, accurate, and complete sentences?
3. Does he enunciate distinctly?
4. Does he have suitable pitch and tonal quality?
5. Does he maintain poise?
6. Does he utilize sufficient variety of words?
7. Does he hold to his subject and his sequence within the subject?
8. Does he present the conclusions from various points of view?
9. Does he state reasons and authorities for his conclusions?

Evaluation by the Pupil

The pupils must share in the evaluative process in the modern secondary classroom. It is a cooperative undertaking to be developed by both teacher and pupil. It is well to understand at the outset that while there are some aims in which the evaluation is shared, there are some other aims that the teacher evaluates without the pupils' aid.

The development of standard and criteria for evaluation and the application of these is definitely part of the learning process for the pupil. Many classes have developed standards for judging oral reports, pictures, exhibits, and numerous other activities. There is a discussion period during which a comparison is made, showing the extent to which the activity approaches the standard. Teachers must be very careful to see that this process of pupil evaluation is impersonal, in terms of the standard rather than in terms of personalities.

A Final Word

Evaluation has a definite place in the learning process which takes place in the modern secondary school classroom. The evaluation should be in terms of the aims of the course or unit which is in progress. Both tests and judgments are utilized in determining to what extent the aims have been realized by individual pupils. Whenever possible the development of these means of evaluation should be a cooperative undertaking of pupil and teacher.

Finally, it should be understood that the principal use which is made of the evaluations is not in the marking of pupils, but in the *improvement of instruction.*[3]

[3] For a comprehensive discussion of the many uses of test results in the classroom see: J. Murray Lee, "A Guide to Measurement in Secondary Schools." New York: D. Appleton-Century Co., to be published in January, 1936.

THE USE OF SCHOLASTIC APTITUDE TEST DATA IN ELEMENTARY SCHOOL EDUCATIONAL ADMINISTRATION

ALFRED S. LEWERENZ
ASSISTANT SUPERVISOR, EDUCATIONAL RESEARCH AND GUIDANCE SECTION
LOS ANGELES CITY SCHOOLS

SCHOLASTIC aptitude test data may be used as a prediction of the academic attainment possible for an entire school as well as for an individual pupil. In the elementary schools of our city batteries of scholastic aptitude and academic achievement tests are given at fairly regular intervals as part of the educational guidance program. The test results for each child are analyzed to determine whether his educational program in the subjects tested is in accord with his present capacity for learning those subjects. In addition the test results are summarized by means of Hollerith equipment to give a picture of each school as a whole. These summaries show the average accomplishment by grades for each subject tested. As part of the report a distribution is also made of the age-grade data to reveal any undue retardation or acceleration that may be present.

Summary data have been accumulated over a period of years and are now available as a basis for preparing administrative guides. Two such guides that are now in use are:

I. A table showing the relation of IQ to chronological age in the classification of pupils.

II. Percentile tables giving the academic achievement of schools of varying IQ.

The first tabulation is useful to a principal in checking his classification to see whether marked chronological acceleration or retardation is present. By noting the intelligence factor it is possible, for example, to determine whether a school with a given amount of chronological retardation is justified in having over-age pupils because it has a population below average in scholastic aptitude. The second set of tables is used in reporting survey summaries to schools. The tables are prepared so that a school may compare itself with other schools of approximately the same intelligence level. The manner in which these tables are used is given in greater detail below.

I.

In the schools of Los Angeles there are a number of provisions for pupils with special educational needs. For beginning B1 pupils there are Transition Rooms or Groups receiving entrants with less than 76 months mental age. Pupils in these rooms are considered of insufficient mental maturity to begin formal reading and their instruction is modified to include experiences which will prepare them for the regular B1 work a semester or so later. Development centers and classes care for marked cases of mental retardation through all the elementary grades. Here hand-work and a simple curriculum emphasize health, conduct, and skill in activities having vocational or life application possi-

bilities. Pupils whose intelligence quotients are above 125 are recommended for placement in Opportunity rooms where enriched instruction is provided. Dull-normal children who probably would not succeed in junior high school in the regular courses may be sent from elementary schools at the age of 14 on special transfer. These pupils are provided with a modified curriculum in the secondary schools. A basic rule guiding promotions is that rarely should children be more than a year over-age or under-age and almost never should be as much as two years chronologically accelerated or retarded.

To knit together the various provisions for special types of pupils and to provide a basis for checking a pupil's classification, a table has been prepared that utilizes the two vari-

LOS ANGELES CITY SCHOOL DISTRICT
DIVISION OF CURRICULUM
Educational Research and Guidance Section

SUGGESTED CLASSIFICATION OF PUPILS BY AGE AND I. Q.
When Academic Achievement and Social Development Warrant Such Placement

INTELLIGENCE QUOTIENTS

Chron. Age in Months	65-69	70-74	75-79	80-84	85-89	90-94	95-99	100-104	105-109	110-114	115-119	120-124	125-129	130-134	135-139
69-71									B1	B1	B1	B1	B1	B1	B1
72-77	Transition B1							B1	B1	B1	B1	A1	A1	B2	B2
78-83							B1	A1	A1	A1	A1	B2	B2	A2	A2
84-89 (7-0)	Development Classes		B1	B1	B1	B1	A1	A1	A1	B2	B2	A2	A2	B3	B3
90-95			B1	B1	B1	A1	B2	B2	B2	A2	A2	B3	B3	A3	A3
96-101			A1	A1	A1	B2	A2	A2	A2	B3	B3	A3	A3	B4	B4
102-107			A1	A1	A1	A2	B3	B3	B3	A3	A3	B4	B4	A4	A4
108-113			B2	B2	B2	B3	A3	A3	A3	B4	B4	A4	A4	B5	B5
114-119			A2	A2	A2	A3	B4	B4	B4	A4	A4	B5	B5	A5	A5
120-125			B3	B3	B3	B4	A4	A4	A4	B5	B5	A5	A5	B6	B6
126-131			A3	A3	A3	A4	B5	B5	B5	A5	A5	B6	B6	A6	A6
132-137			B4	B4	B4	B5	A5	A5	A5	B6	B6	A6	A6	B7	B7
138-143			A4	A4	A4	A5	B6	B6	B6	A6	A6	B7	B7	A7	A7
144-149			B5	B5	B5	B6	A6	A6	A6	B7	B7	A7	A7	B8	B8
150-155			A5	A5	A5	A6	B7	B7	B7	A7	A7	B8	B8	A8	A8
156-161			B6	B6	B6	B7	A7	A7	A7	B8	B8	A8	A8		
162-167			A6	A6	A6	A7	B8	B8	B8	A8	A8	Senior			
168-173 (14-0)			Spec. Trans.			B8	A8	A8	A8	High School					

NOTES

1. Educationally mal-adjusted cases should be considered as candidates for opportunity and adjustment rooms.
2. Pupils below 75 I. Q. may be considered candidates for development centers or rooms.
3. Children falling within the dotted rectangle on the right are probable candidates for opportunity A rooms.

Approved: Elizabeth L. Woods, Supervisor
Arthur Gould, Deputy Superintendent

Revised 5-12-1932

ables, chronological age and I Q. The basis for the grade placements is the actual age-grade distribution in Los Angeles. Pupils within the range of normal intelligence are given grade placements in agreement with existing conditions. Dull and bright children are given sufficient retardation and acceleration to tie in with existing procedures without at the same time violating the educational philosophy which attempts to prevent pupils from getting too far removed from their own social groups.

On the basis of the accompanying table the following recommendations are suggested for the types described above:

Chronological Age in Months	*IQ*	*Placement*
72	80	Transition B1
84	80	Regular B1
96	70	Development class
72	100	Regular B1
114	110	A4
102	130	Opportunity A (highly endowed)
168	80	Special transfer to junior high school

Pupils who are of normal intelligence but are educationally maladjusted are recommended for remedial instruction in an Adjustment room.

The table is not intended as a basis for promotion. It is used only in conjunction with facts relative to a given pupil's academic achievement, health, and social development. It is an aid to principals and counselors in:

A. Correcting undue under- or over-ageness by discovering the pupils who create the condition.
B. Selecting pupils who are candidates for special rooms.
C. Ascertaining which pupils are to be promoted to junior high school on special transfer.
D. Determining the grade in which a new child entering from outside of the district should be placed.
E. Providing standards for use in guiding promotion policies.

When, for example, a school's survey summary shows that the A6 grade is definitely under-age and that academic accomplishment is below grade, the table is of material assistance in locating the pupils who are creating the condition. In such a situation the cause often may be attributed to too rapid promotion of pupils of less than average ability. A child, 133 months old and with an I Q of 96 would probably find some difficulty in doing average work in a regular A6 class. He would be better placed in an A5 group according to the table.

It is also often the case that children of high scholastic ability have not been so recognized by their teachers. As a result these brilliant pupils are really only loafing along as no attempt is made to enlarge their educational experiences. Their educational achievement accordingly will remain near the average of the group. As a case in point, we can take the example of a pupil 115 months old with an I Q of 126 who has not been accelerated so that he is in the low fourth grade along with children of average ability. Experience has shown that this same pupil could successfully accomplish the work in a B5 class as is shown on the chart. It is also indicated that he is a candidate for a room for children of superior mentality. At the earliest opportunity he would be given an individual mental examination, such as a Binet, to verify his scholastic aptitude. If placed

LOS ANGELES CITY SCHOOL DISTRICT
DIVISION OF CURRICULUM

Educational Research and Guidance Section

Percentile Distribution of Survey Data for Schools I. Q. 85-89

The survey made at.....School during the month of.....193.... gave averages as indicated below. This chart compares the school named with other schools of the same intelligence level. Survey data above the median (50th percentile) indicate that the school is better than the average school of the same intelligence level while figures below the median reveal a standing lower than average.

Percentile Rank	Av. Read. Gr. Pl.	Av. Vocab. Gr. Pl.	Av. Fund. Gr. Pl.	Av. Reas. Gr. Pl.	Av. Chron. Gr. Pl.	Av. Intel. Gr. Pl.	
100 High	—.10	+.20	+.65	+.10	+.52	—.08	
90	—.20	—.05	+.35	—.10	+.60	—.17	
80	—.30	—.40	+.07	—.25	+.65	—.23	
70	—.50	—.50	—.20	—.45	+.70	—.37	
60	—.60	—.53	—.36	—.61	+.73	—.53	
50 Av.....	—.70....	—.57....	—.38....	—.63....	+.81.....	—.64	Median
40	—1.10	—.60	—.55	—.70	+.89	—.72	
30	—1.50	—1.10	—.72	—.78	+.99	—.75	
20	—1.80	—1.38	—.77	—1.00	+1.12	—.85	
10	—2.30	—1.43	—.81	—1.85	+1.33	—1.02	
0 Low	—2.70	—1.50	—.88	—3.00	+1.65	—1.20	

Approved:
Elizabeth L. Woods, Supervisor
Arthur Gould, Deputy Superintendent

in an Opportunity room he would be given ample opportunity to develop his special abilities with more freedom than is found in the usual classroom.

II.

The tables giving academic achievement of schools with varying I Q, arose from a need for providing fairer standards of comparison when interpreting survey results. The average I Q of schools in the Los Angeles system varies from 82 to 114. This spread means that some schools, on the average, are within the "dull-normal" classification while others are in the "superior intelligence" category. Such a range in potential ability necessarily is reflected in about the same range of educational accomplishment. It is obviously unfair to penalize a school markedly below the average in scholastic aptitude if achievement test results are "not up to norm." It is also true that another school whose survey results are "way above norm" may be actually doing comparatively poor work because of the very superior type of pupil in that school.

With the purpose of determining what accomplishment might reasonably be expected of a school in the light of its average I Q, a study has been made of all the recent survey records available in Los Angeles. A total of more than 270 schools is represented. Table I gives the number of schools represented for each item at each I Q level. Table II gives the 1931-32 tenth month enrollment for the schools represented in the data for

reading. These figures are representative of other items. From the accumulated data six percentile tables were prepared. These tables represent five-point intervals of I Q, ranging from 85 to 114, two of which only are included, that for 85-89 I Q, and that for 110-114 I Q.

TABLE I

Number of Schools From Which Data Were Available

IQ Level	*Read.*	*Vocab.*	*A. Fund*
85-89	10	9	9
90-94	25	25	25
95-99	39	38	39
100-104	86	84	88
105-109	84	84	85
110-114	25	26	28

IQ Level	*A. Reas.*	*Chron. Gr. Pl.*	*Intell. Gr. Pl.*
85-89	9	10	10
90-94	25	25	25
95-99	39	38	37
100-104	84	80	81
105-109	85	78	82
110-114	26	27	27

TABLE II

Pupils Enrolled in Schools Represented in the Distribution for Reading by I Q Level of Schools

IQ Level	*Enrollment*
85-89	4,269
90-94	13,424
95-99	22,212
100-104	41,598
105-109	45,169
110-114	13,707
Total	140,379

When a testing program has been completed by a school and the results have been summarized by the Educational Research and Guidance Section the data are indicated on the percentile distribution appropriate to the school's average I Q. In this way it is possible for a school to note how it stands not only with regard to "national" norms but also how it compares with other schools in the city having pupils of about the same average mental level.

Let us look at the first percentile table which is for schools of 85-89 average I Q. A glance at Tables I and II shows that there are about ten schools represented with some 4,269 pupils furnishing the data for the percentiles. The entries are all in terms of fractions of a school year above or below norm as is usually determined by a standardized test. The school year is considered as having ten school months. To illustrate what is meant, let us say that a low fourth-grade class in the third month of school receives an average grade placement of 3.1 in reading comprehension. The grade norm is 4.2 so the class is one year and one month below grade norm. This deviation below norm is expressed in figures as —1.1.

Coming back again to the table for schools of I Q 85-89 it will be seen that subject accomplishment is for the most part below norm even for the best schools in the group.

The average situation is as follows:

A. Reading comprehension is seven months below norm.
B. Vocabulary or word recognition is more than five months below national norms.
C. Arithmetic fundamentals is almost four months behind.
D. Arithmetic reasoning is more than six months retarded.
E. Chronological age is eight months retarded.
F. Intelligence or mental age is more than six months below norm.

Schools with grade placements above the average for their group may be said to be working above expectation while the opposite is the case with schools placing below the 50 per-

LOS ANGELES CITY SCHOOL DISTRICT
DIVISION OF CURRICULUM
Educational Research and Guidance Section
Percentile Distribution of Survey Data for Schools I. Q. 110-114

The survey made at.....School during the month of.....193.... gave averages as indicated below. This chart compares the school named with other schools of the same intelligence level. Survey data above the median (50th percentile) indicate that the school is better than the average school of the same intelligence level while figures below the median reveal a standing lower than average.

Percentile Rank	Av. Read. Gr. Pl.	Av. Vocab. Gr. Pl.	Av. Fund. Gr. Pl.	Av. Reas. Gr. Pl.	Av. Chron. Gr. Pl.	Av. Intel. Gr. Pl.	
100 High	+1.30	+1.20	+1.10	+.95	—.53	+1.00	
90	+1.10	+1.01	+.97	+.63	—.32	+.97	
80	+.85	+.91	+.81	+.54	—.28	+.90	
70	+.79	+.67	+.74	+.40	—.24	+.84	
60	+.70	+.56	+.63	+.28	—.21	+.81	
50 Av.....	+.64....	+.42....	+.57....	+.18....	—.20....	+.77	Median
40	+.57	+.37	+.45	+.09	—.18	+.73	
30	+.53	+.31	+.34	+.06	—.17	+.70	
20	+.40	+.23	+.16	—.08	—.15	+.66	
10	+.17	+.13	+.03	—.20	—.06	+.64	
0 Low	+.05	—.15	—.20	—.32	+.15	+.50	

Approved:
Elizabeth L. Woods, Supervisor
Arthur Gould, Deputy Superintendent

centile point. Of course, no one school is represented by the row of figures at the 100 percentile point. The percentiles for each column were worked out separately. It would, therefore, be very unusual for a school to be near the 0 or 100 percentile row for all headings.

Results from a school survey summary are graphed in on the appropriate blank with a red pencil and sent to the school principal. Copies are sent to the assistant superintendent in charge and the Course of Study Section.

Let us say that a school with an average I Q of 86 received the following averages:

Reading comprehension	—.90
Vocabulary (word recognition)	—.70
Arithmetic fundamentals	—.63
Arithmetic reasoning	—.72
Chronological age	+.56
Intelligence (mental age)	—.76

If the above figures were chartered, the percentile ranks would be somewhat as follows:

Reading comprehension	—.45th percentile
Vocabulary (word recognition)	—.38th percentile
Arithmetic fundamentals	—.35th percentile
Arithmetic reasoning	—.37th percentile
Chronological age	+.95th percentile
Intelligence (mental age)	—.29th percentile

The school's accomplishment would appear below the average for other schools of the same type. Examination of the data for chronological and mental age gives a possible answer as to the cause of the apparent deficiency. It will be seen that the pupils are two and a half months younger than the average chronologically and one and two-tenths months younger mentally than the average for schools in the same I Q group.

This chronological and mental acceleration is great enough to explain why the pupils are below expectation in their subjects. In all probability, the pupils are working up to their capacity but an error in classification, such as too rapid promotion, has put the pupils in a grade above that usual for their degree of mental maturity. It is in a case like this that the table described in the first part of the article is useful in determining just who are the children unduly accelerated. By comparing each child's actual grade placement with that recommended on the chart it is possible to make up a list of pupils to be held for non-promotion. A case of an opposite nature to the foregoing is also typical. A school of 111 average I Q received the following averages for a survey:

Reading comprehension	+.55
Vocabulary (word recognition)	+.39
Arithmetic fundamentals	+.57
Arithmetic reasoning	+.20
Chronological age	—.18
Intelligence (mental age)	+.82

Reference to the chart for I Q 110-114 gives percentiles as follows:

Reading comprehension	35th percentile
Vocabulary (word recognition)	45th percentile
Arithmetic fundamentals	50th percentile
Arithmetic reasoning	52nd percentile
Chronological age	40th percentile
Intelligence (mental age)	63rd percentile

In this example a school of high average intelligence appears to be well above norm in reading and word recognition until compared to other schools of the same level of ability. To make matters worse, the pupils are distinctly older than those in other similar schools, both chronologically and mentally. There is evidence that greater attention must be given to reading methods and that the principal should eliminate the over-ageness present as rapidly as possible.

The validity of the tables is a question which is important and which perhaps is answered in the following paragraphs.

The largest group of schools for which data are available is in the 100-104 I Q group. Between 80 and 86 schools contributed their averages to the percentile columns. Test results from approximately 40,000 pupils were included in the surveys. Consequently the percentiles for this group should be fairly reliable. While there are several fewer schools in the next higher group, that of 105-109 average I Q, there are about 4,000 more pupils included.

The percentile table for the I Q group of 100-104 probably includes what most nearly corresponds to the average schools in the system inasmuch as the average I Q of pupils in regular elementary schools is known to be near 104.

It will be seen that the 50th percentile point for this 100-104 group is slightly above norm except in arithmetic reasoning where it is a very small amount below norm, about six days if the —.03 were to be interpreted in terms of time. The consistency of these averages is one verification of the validity of the table. Another check is that when the 50 percentile points for all tables are charted on any one table they are found to be mutually exclusive and to travel consistent independent paths when graphed. Actual experience for more than a year has shown that the tables are a practical means of interpreting survey summaries.

In conclusion it might be pointed out that:

A. The chronological grade placement column has the increments running in the opposite direction from the others as over-ageness is regarded as being undesirable.

B. The columns for chronological and mental age are used to interpret more exactly the significance of the subject test results.

C. Achievement may appear high on the chart but a glance at the chronological column may show the school to be markedly over-age. On the other hand, poor subject accomplishment may be attributable to the children being mentally immature.

D. When a school is low in subject accomplishment and is at the same time markedly over-age, there is a clear indication that the situation warrants investigation.

E. Schools which are above expectation in the subject tests and which at the same time show a small amount of chronological acceleration are usually considered as doing excellent work.

RECENT DEVELOPMENTS IN EDUCATIONAL MEASUREMENT

DAVID SEGEL
SPECIALIST IN TESTS AND MEASUREMENTS
UNITED STATES OFFICE OF EDUCATION

THE changing emphasis in education has naturally influenced the direction of work in the measurement field. Whereas the first ventures in testing — such as the survey of achievement in the traditional fundamental subjects of the elementary school and the measurement of general scholastic aptitude (intelligence) — were for long considered quite adequate and all inclusive in themselves, there now seems to be a tendency to place the emphasis rather on social outcomes and to search for particular aptitudes in individuals. That is, testing with some pre-conceived standard of intellectual attainment in mind, in order to work toward achievement of such a standard, has been replaced to some extent by an attempt to measure the development of the whole child, not only as to intellect, but also in such phases as appreciation, attitudes and social conduct.

Two aspects are apparent in this new movement. One is the emphasis on individual development — i.e., that the individual should develop as many different capacities as possible and especially those which appear to give the most promise for his future growth. The other aspect is an attempt to so educate that the child will see the proper relationship between the various elements in life. This involves the integration of subject matter on the one hand and a definite attempt to socialize the individual on the other hand.

Various lines of measurement are being developed to satisfy this new approach in education. In connection with the cultivation of a larger number of areas of skills, traits, and interests in the individual, certain lines

of progress can be seen. First may be mentioned the effort to measure all the objectives in a single subject, i.e., a program of testing in a particular subject which includes all objectives and is not merely a check of facts memorized. Too often tests have been hastily thrown together without regard for the different objectives of a subject. They have been constructed to test the recall of the bare facts of a subject, but issued as though comprehensive tests of the subject. The work of reconstructing our achievement tests to cover all the outcomes of instruction in a subject is perhaps best shown by the work of Tyler[1] at Ohio State University. Tyler has cooperated with instructors of different college subjects in analyzing their subjects and in setting up the objectives in such definite terms that examination items may be easily constructed. For example, for college zoology it was determined that the major objectives to be examined were:

(1) The ability to recall important zoological facts.

(2) The ability to remember and state zoological principles.

(3) The ability to formulate an appropriate definition or description of the common technical terms in zoology.

(4) The ability to formulate in his own words as complete a generalization as is justified by the data presented, when the student is given the results of an experiment.

(5) The ability to determine what needs to be established in order to substantiate a given hypothesis, and the ability to plan an experiment to establish or disprove these facts.

(6) The ability to predict the outcome of a situation involving one or more zoological principles included in the course and the ability to use these principles in explaining why the outcome could be expected.

(7) Skill in certain significant laboratory techniques.

(8) The ability to report the results of experiments in effective English.

(9) The ability to state the sources which were most likely to give dependable information on specific kinds of zoological problems.

Another line of effort is the attempt to find new ways of measuring social traits. The tests which have generally been used for such measurement are performance and opinion tests. It has been found difficult with these tests to conceal the real purpose of the testing from the pupils, and the results have, therefore, not been satisfactory. A method which obviates most of this difficulty is the word association test by Kelley[2] in a study made as a part of the investigation of the social studies. In a word association test there are so many words that are possible as response words that the person making the responses cannot have much of an idea as to the significance of the different responses. The test as originally constructed has some validity in results with groups, but it must be refined before it can be used to diagnose the strong or weak traits of individual pupils.

A third line of endeavor is one related to the type of work carried on by Kelley. It is the study of independent traits. Thurstone[3] and Hotelling[4], following the pioneer work of Kelley and Spearman, have been re-

1 Tyler, A. W.: "Formulating Objectives for Tests." Educational Research Bulletin, 12: 197-206 (Oct. 11, 1933).

2 Kelley, T. L., and Krey, A. C.: "Tests and Measurements in the Social Sciences." Scribner's, 1934.

3 Thurstone, L. L.: "Multiple Factor Analysis." Psychological Review, 38:406-27.

4 Hotelling, H. "Analysis of a Complex of Statistical Variables into Principal Components." Journal of Educational Psychology, 24:417-41; 498-520 (Sept. and Oct., 1933).

cently perfecting mathematical methods which will isolate various human traits or tendencies, both social and intellectual. These mathematical principles are being applied to test data to find the various traits. It is clear that this work will in the future lead to better educational and vocational guidance and better instructional practices. The extent of the practical application of this work will depend upon the ingenuity and effort shown in the construction of tests which will faithfully represent the independent traits or tendencies discovered.

Trends in measurement which are related to the second aspect of the growth in educational theory and practice—the trend towards integration of the pupil's whole personality — are clearly discernible. The activity movement and the integration of subject matter has made the measurement of the total behavior of both pupils and teacher in the classroom a matter of concern. There seem to be two ways, at least, in which this problem is being attacked. One method of measuring the totality of instruction is through observing the techniques of instruction and the behavior of the children rather than through definite testing of individual children. A good illustration of the use of such measurement is the work being carried on by Wrightstone[5] in comparing practices in different school systems. In his initial study he rated school efficiency on, among other things, the type of pupil responses, planning and discussion and the teacher conduct of class discussion situations. The rating was done fairly objectively by making classroom observations on set time schedules. The pupil responses observed were:

(1) Initiative shown in prepared voluntary reports or exhibits.

(2) Initiative shown in extemporaneous contributions from real experience.

(3) Initiative shown in extemporaneous contributions from vicarious experience.

(4) Initiative shown in suggesting means, methods, activities, and solutions.

(5) Responsibility shown in preparing assigned reports or exhibits.

(6) Curiosity shown in asking questions on the topic, unit or problem.

(7) Criticism made of a contribution.

The teacher conduct of class discussions observed were:

(1) Allowing pupils to make a voluntary contribution.

(2) Encouraging pupils to make a contribution.

(3) Proposing questions or theses for pupils or class.

(4) Referring pupils to sources of data or information.

(5) Suggesting (explaining) means, methods, activities or solutions.

(6) Discouraging or prohibiting a pupil from making a contribution (a negative phase).

(7) Recalling a pupil's attention by direct word, look or gesture, etc.

This method is not dependent upon any particular subject matter which is being taught but measures the school on the basis of what is thought to be the type of activity which will develop the pupils into integrated personalities. Such methods have not been applied, except in a few instances, for the purpose of rating individual pupils. At present they are useful mainly in measuring classes or schools.

The first trend in testing discussed in this paper is the trend towards testing all the outcomes in a *definite sub-*

[5] Wrightstone, J. W.: "Appraisal of Newer Practices in Selected Public Schools." Bureau of Publications, Teachers College, Columbia University, New York City, 1935.

ject field. There exists also a broad field in that of testing a unit of work which may include elements from *many different restricted subject* fields and more generalized outcomes than can be hoped for from any one subject. This development is important because it is one which immediately affects the instructional process in that it helps the teacher analyze and classify the subject matter to be taught and makes known immediately whether or not the pupils have achieved. Such tests are more likely to be constructed by the teacher than other types of tests and, therefore, are more likely to force an analysis of the learning situations in the units. Even when not constructed by the teacher, they are considered important by her because they are so intimately connected with her immediate work. Of course, only some unit tests cover integrated work. Many cover a restricted subject field. This is logical where the instructional unit comprises some one skill. Where an instructional unit is built around a topic or several related topics a somewhat new approach in testing is required. For example, in measuring the outcomes of a unit built around the life of the American Indian in the fourth grade of the elementary school the check-up would cover not only historical facts but also geographical items, English composition and certain social attitudes of the people of other races.

A plan for the construction of tests for one school system[6] shows the possibilities of the use of tests in units of instruction. The plan is briefly:

1. Diagnostic tests for each of the activities which are suggested for a given unit. This, of course, will involve a careful study of the materials, as well as the objectives. It will make teachers aware specifically of what they are trying to do. These diagnostic tests will then be used in two ways. Where the conditions warrant, the teachers may use them as pre-tests before the activity gets under way. But as each activity draws to a close, the teacher may use them for diagnostic purposes in order to determine specifically what outcomes have not been reached and what additional developmental or drill work is needed.

2. After the various individual activities have been concluded and special culminating activities, if any, have been completed, a more comprehensive standard test of the unit may be given. These standard tests will consist of the essential specific outcomes in the way of facts, information, principles, problems, skills, appreciations, ideals and attitudes, so far as we are able to determine them.

The new approaches in testing which have been discussed in this paper are consistent with the added socialized aims in learning and to some extent these new techniques have been developed because of the changing emphasis in education. A study of these more recent practices in educational measurements does not imply any lessening of interest in or use of practices already established. Development of these newer procedures, however, is evidence of the expanding use of the scientific method in education. The writer feels confident that educational psychologists will be able to measure practically all of the outcomes of instruction which are now or may in the future be legitimately claimed by educators.

6 Burbank City Schools, Burbank, Calif.

THE CONSTRUCTION AND STANDARDIZATION OF THE KNAUBER ART TESTS

ALMA JORDAN KNAUBER
SCHOOL OF HOUSEHOLD ADMINISTRATION
UNIVERSITY OF CINCINNATI

Purpose

IN a few of our larger, more progressive cities, opportunities are being provided for the artistically gifted, but standardized and objective ways of discovering these students and measuring the amount of their ability have not been used. The judgment of the grade or art teacher was depended upon to grant the opportunities, and experience has shown that this procedure has not always been reliable. There are many studying art who would be better fitted for some other line of work. In an article in the *New York Times* of Sunday, July 28, 1935, Forbes Watson, in commenting on the art work done through P. W. A., said: "I have examined the work of at least three thousand artists. Of these, I should say that about five hundred could fairly be classified as artists without quotation marks. . . . " Just as there are many such self-styled "artists" who in reality belong in another field of work, so too there exist some potential creative artists who are never discovered. This situation results in an economic and social loss. Every one should be given an opportunity to study the appreciation of art and to learn its uses in everyday life. Every one who desires to may study art as an outlet for his creative urge and as a hobby, or to help him appreciate; but only those who are specially gifted should study art from a professional veiwpoint. Those who possess art-ability should be found early and given the best training possible, for they are of inestimable value to a country. Every student of art should know of his potential success in the field before he spends time and money in preparation for it.

Construction

The Knauber Art-Ability Test was devised to enable the teacher or director of art to obtain a definite and objective measure of different degrees of artistic talent. The first form of this test was printed in 1926, and the test has been revised and printed twice since then. The second form of the test was presented, upon invitation, to the Carnegie Committee for Research in Art. At this meeting, it was suggested that the test be made into three individual tests: an art-ability, an art-vocabulary, and an art-appreciation or art-judgment test. This was done, and the art-ability and art-vocabulary tests have been printed and are now complete, with an examiner's manual for each test.

The Knauber Art-Ability Test is a performance test which gives the student ample opportunity for creative expression. The seventeen problems are scaled from simple to more difficult ones. They necessitate the testing of short-time memory, long-time memory, observation, accuracy, ingenuity, creative ability, critical faculty, and so on. The last problems, which

are very difficult, are valuable for measuring the student's ability to handle the most advanced forms of art, such as, for example, symbolism and expressionism. The first two pages of questions, which are not to be scored, give the examiner a chance to know the interests, ambitions, and background of the student.

The Knauber Art-Vocabulary Test is a test of art-knowledge. It consists of one hundred multiple-choice art terms which are arranged in order of their difficulty.

Standardization

To secure a wide sampling of scores for standardization purposes, test scores were obtained from schools in all parts of the United States — from New York to Los Angeles, and from Chicago to Austin. The junior and senior high school grades were from the public schools of Ashtabula, Ohio; Cincinnati, Ohio; Long Beach, California; Los Angeles, California, and Washington, D. C. The colleges which participated were: Art Institute of Chicago, Columbia University, Edinboro State Teachers College, Johns Hopkins University, Kanawha College, Miami University, Ohio Mechanics Institute, Pratt Institute, Rhode Island School of Design, University of Chicago, University of Cincinnati, and the University of Texas. These participants included students of art, education, and psychology. The grades tested were seven through sixteen.

Development of Scoring Key

To standardize the art-ability test 1366 cases were used. To standardize the art-vocabulary test 1907 cases were used. These scores were distributed among ten different grades. All the tests were scored by the author. During the scoring process, a scoring key for the art-ability test was evolved. First, the problems were marked horizontally, that is, the first problem of all the tests, then the second problem, and so on, to the end of the test. The scores given were checked. Three levels as a guide for scoring were decided upon, namely, ten or excellent, six or average, and three or poor. Each problem was rated 0-10. During the scoring process, examples of the three levels, ten, six, and three, were taken from the students' papers. To make the scoring key more accurate, this selection of examples which were scored by the author, was submitted to three judges: an artist, a commercial artist, and a landscape architect. Each rated the 270 examples, and their estimates were tabulated. There was a 73% agreement with the author's scores. From those examples on which the judges agreed 160 were used for the scoring key. In order to make the scoring key as objective and accurate as possible, and to aid the scorer, a statement which supplements the examples in the three levels of each problem is given.

Reliability of Scoring Key

To test the reliability of the scoring key, five college students, namely, two freshmen, two sophomores, and one senior art student, scored a set of eight tests which had been scored and selected by the author. The tests were chosen to include various degrees of ability so as to secure from the students as wide a range of scoring as possible. The students were not told about the method of selecting the tests. These tests presented forty-

eight total scores or 816 problem estimates, for each test contained seventeen problems. A comparison between the total scores given by the author and those given by the college students was made, and a coefficient correlation of +.97 was obtained. This correlation shows that the scoring is reliable when done by average college students. It also shows the high degree of objectivity in the character of the problems in the test. This indicates not merely a close correlation between the scoring by the author and students, but also a close similarity of scoring among the students. Although there were only eight cases in this group, the correlation coefficient of +.97 indicates adequate reliability. Pearson's Product Moment Formula was used.

Norms

After all the art-ability and art-vocabulary tests were scored, the scores for each grade were tabulated in numerical order from the highest to the lowest. The middle score of each grade was taken for the norm of that grade. The aim was to have at least one hundred tests for each grade before establishing a norm, but in a few instances the number exceeded one hundred, and in a few cases, the number was less than one hundred. When the norms were checked and it was found that the norm was either too high or too low for its position in the scale; because of too wide a sampling of art or non-art majors it was necessary to interpolate between the norms above and below.

Groups

The levels of art-ability were chosen to parallel the grading system A, B, C, D, and E, used in some schools, or its equivalent, excellent or exceptional, good, average, poor, and very poor. These levels were worked out in several ways: first, the upper sixth of the cases in each grade were taken as exceptional, one sixth as good, two sixths as average, one sixth as poor, and one sixth as very poor; then the levels were worked out by using a percentage basis of a normal probability curve, 7%, 24%, 38%, 24% and 7%. An average was taken between these two methods when the levels did not coincide. The levels were also rechecked by reference to a number of students' papers; for example, all the tests which fell in the exceptional level were questioned as to whether or not they displayed exceptional art-ability; and as a result several of the levels were changed.

Reliability of Art-Ability Test

The reliability was found by dividing the tests of eighty-three students from various grades into two parts, with the same number and similar problems in each part. The first part included Problems 1, 4, 6, 8 (first half), 11, 13, 9, 15, 16, 3. The second part included Problems 2, 5, 7, 8 (second half), 12, 14, 10, 15, 17, 3. The two parts of Problem 8 were graded separately for this experiment in order to make the halves of the test equal in number of problems. Since there were no problems to balance numbers 3 and 15, both were used in each half. The total scores made on each part were compared, and a correlation of .95 .007 was obtained. This correlation indicates a high degree of reliability. The exceedingly high correlation might be

interpretated as indicating that one part of the test would be sufficient, but the larger sampling of a student's work tends to make the test more accurate as a measure of his art-ability.

Nine college students who happened to be in classes tested in two successive years, took the test each year. These students were not art majors, but were taking a required art subject. Although the number of cases was small, the group included the several grades and various degrees of art-ability. The two scores made by each of the students were compared and a coefficient of correlation of +.96 was obtained. This correlation, even after a year's interval, indicates a high degree of reliability. Those who scored high on the first test also scored high on the re-take, and those with a low score the first time, made almost the same scores the second time. This also suggests that the test is a measure of native art-ability rather than of acquired art skills.

Validity of Art-Ability Test

By validity we mean the extent to which a test measures that which it is supposed to measure. In order to establish the validity of a given measure, it is necessary to compare the ratings resulting from this measure, with ratings resulting from some other established criterion of the ability in question; and to check the contents of the measure (test) with an accepted criterion.

No satisfactory criterion for statistically determining the validity of the art-ability test is available. The art grades given by the teacher are not objective and do not estimate *only* the native art-ability of the students; so these grades would not be a satisfactory basis of comparison to determine the validity of the test. No other measure, testing only native art-ability, is available to which this test could be compared. The test has been submitted to authorities in the field of testing and of art, and subjected to practical use by the author in her art classes. Comparisons have been made to check the test in various ways.

A comparison was made between the scores obtained by college students who were majoring in art, and the scores of college students who were not art majors. The medians of both groups were found and compared. The median for art majors of the junior class in college was 95, and for non-art majors of the same class it was 52. Thus it may be seen that the medians for art majors were considerably higher than for the non-art majors. From this it follows that the test meets a common-sense criterion for estimating the amount of native art-ability possessed by a student.

A study was made of the scores of a group of art teachers, who were doing graduate work in art, and a group of university graduates who were taking an art course but were not art majors. The latter group was doing graduate or fifth-year work in teacher's training. The median for the art teachers was 123, and the median for the non-art teachers was 61. This study shows another value of the test in measuring art-ability.

Reliability of Art-Vocabulary Test

To check the reliability of the art-vocabulary test, sixteen college students who happened to be in classes tested in two successive years took the

vocabulary test each year. Although the number of cases is small, they are representative of a larger group, because it is a natural sampling from each college grade. The two scores made by each of these students were compared, and a coefficient of correlation of $+.67 \pm .093$ was obtained. This correlation is not high; but since this is a test of acquired knowledge, the interval of a year was too long. From a study of the scores it was observed that those who scored lowest the first time, made consistently low scores the second time, and that the students with a larger art-vocabulary tended to improve it through interest or study during the year.

The reliability of the art-vocabulary test was found by dividing the tests of 100 students, of various grades, into two parts, odd and even questions. The total scores made on each part were compared, and a coefficient of correlation of $+.78 \pm .03$ was obtained. Although this is not high, it merely indicates that the questions are not all arranged in their order of difficulty. The scaling was done on the basis of scores obtained only from local students who had had similar training. A higher correlation between the halves might have resulted if the scores for this purpose had been obtained from students with different training and from various parts of the United States.

Validity of the Art-Vocabulary Test

Since there were no satisfactory criterion for statistically determining the validity of the art-vocabulary test, comparisons were made to check the test in various ways.

Scores of art and non-art students were compared. The scores were from 84 art majors and from 90 non-art majors composed of graduates, seniors, and juniors, in college. The median for each of these groups was obtained, and it was found that in each group the median was consistently lower for the non-art students. The ratios of the medians in each of the three groups was approximately the same. The above findings suggest that this test is capable of measuring art-knowledge, because it shows that the art-vocabulary of art students is consistently higher than the art-vocabulary of the non-art students.

To analyze the test further, a study of the tabulations of students' scores was made and it was found that there is a gradual rise in the highest scores made in each grade: beginning with the seventh grade they are 62, 53, 63, 64, 66, 74, 74, 81, 89, 92. The medians for these same grades are 19, 23, 25, 36, 37, 39, 44, 54, 64, 66. The steady rise of the highest scores made indicates that this is a test of art-knowledge and not one of native art-ability. Although this rise is not perfect, owing to variations in the grades tested, it is in marked contrast to the highest scores made on the art-ability test. The highest scores made on the art-ability tests are 48, 107, 122, 127, 131, 114, 131, 137, 143, 137. These do not show a progressive rise from the seventh grade through college.

These comparisons and correlations offer obvious proof of the reliability of the tests. Furthermore, they have been used by the author with marked success to aid individual students in adjusting themselves; for instance,

through the medium of these tests, a college senior majoring in English, but taking freshmen art as an elective, was found to have artistic talent. On the strength of this discovery, she elected more art courses, took advanced work the following year, and was awarded an M. A. degree in art. She taught art in a college and has since spent a year abroad studying art. In the field of art she has made a valuable contribution which would have been improbable without some such means of self discovery as these tests provided. A more systematic follow-up of students who have in the past taken the test and have verified the findings of the tests would increase their already certain reliability. This study suggests that these tests meet the practical requirements needed to prognosticate success in art and to determine the amount of a student's art knowledge.

SOME PROBLEMS IN MEASUREMENT

ERNEST O. MELBY
DEAN, SCHOOL OF EDUCATION, NORTHWESTERN UNIVERSITY

SPEAKING roughly we can now look back upon a quarter century of effort at exact measurement in education. Practical school men and women who have followed the movement from its early beginnings no doubt differ greatly in their appraisal of the amount of progress which has been made. It is probably true that the bitterest critics of the measurement movement tend to underestimate its contributions to education. When, however, one considers the progress which has been made in the light of the hopes of twenty years ago, even those most enthusiastic about exact measurement will feel considerable disappointment. This more or less general disappointment with the results of our measurement efforts is a condition which is likely to retard progress. On every hand one notices a skepticism in regard to tests. Teachers are giving less effort to the construction of testing instruments, chiefly because they doubt the wisdom of the whole testing movement. We may thus face a period in which there will be a reaction against testing and in which there may be little progress in the solution of measurement problems. It is the purpose of the present article to outline a few of the practical problems which are emerging in measurement. The problems are discussed from the point of view of the teacher or administrator rather than from the viewpoint of the measurement expert.

It seems paradoxical that the very period during which the measurement movement has developed should also be the period during which our educational philosophy has developed along lines which make our present tests less and less useful as devices for measuring the outcomes of education. In general, the whole measurement movement has centered around a subject-matter concept of education. Educational growth has been measured in terms of the degree of mastery in specific subject-matter skills and knowledges. Rarely has any attention been given to the experiences of children in acquiring the skills or

knowledges they possess. What is equally important, the measures used threw no real light upon the reasons for failure to acquire the skills tested. Finally, our measurements gave us little or no information on the appropriateness of the subject-matter to the needs or abilities of the pupil. To summarize, we may say that our current tests attempt to measure the child's mastery of certain skills and knowledges without telling us how he secured them or how and why he failed, if fail he did. What is even worse, we measure achievement in a subject without even knowing whether the pupil should have studied the subject at all or not.

Let us now contrast the philosophy basic to our measurements with that which seems to underlie the progressive school. The test starts with subject-matter while school starts with the child. The test stresses results in skills, the school emphasizes personal development. The test places a premium upon what is learned, while the school attaches most importance to learning experiences. The test seeks information about what a child is doing in arithmetic, while school wants to know to what extent arithmetic can serve the child. The test tells us what the child knows, while school wants to know how he lives.

American teachers and parents are slowly realizing the fact that we have in America given the world a new educational philosophy. Our point of departure is that every individual is to be educated. It matters not what his characteristics may be. He may be bright or dull, weak or strong, a future laborer or statesman. The responsibility of the school is to find the activities and experiences by means of which he may make the greatest growth. No subject has any claim under this philosophy except as it can demonstrate its usefulness to the educational growth of an individual child. It is this new educational philosophy which has played havoc with the measurement techniques.

Our first problem then is to develop a program of measurement which is in harmony with our educational philosophy. Modern education seeks a knowledge of the child, such knowledge to be used as a basis for teaching. It is clear that child study is the basic element in the new measurement program. If we search for techniques of child study among our present measurement devices, certain very marked weaknesses of our present procedures will at once become apparent.

Let us take intelligence testing as an example. Specialists in measurement appear to be in constant search for a single measure of intelligence. Some appear to be determined to prove that intelligence is a general factor. Now from the viewpoint of child study, a general measure of intelligence is well nigh useless. The teacher asks, "What kind of intelligence?" From the viewpoint of the teacher a test of musical intelligence, of æsthetic intelligence, of literary intelligence, of mathematical intelligence, etc., would be of inestimable value. For the most part, however, psychologists refuse to work on the development of specific measures of intelligence. In the main they are immersed in a search for a single measure of general intelligence.

Another illustration is to be found

in the constant application of the criteria of objectivity, reliability and validity. Through the worship of this triumvirate of criteria we have fallen into a futile do-nothing formalism. The author was told by a school principal that he would like to do something about character education, but since there were no measures of character development, he thought it best not to attempt character education. It is a curious paradox that education, once a very subjective enterprise, should in a generation have bowed so low in its worship of objectivity. In this respect we seem to differ markedly from other professions. Medicine for example does not hesitate to utilize subjective means, depending for their success upon professionally trained intelligence. Thus an X-ray photograph of the teeth does not supply objective evidence. In the hands of the untrained layman it means nothing. When viewed by a trained specialist it may be of greatest importance. In short, medicine utilizes evidence submitted by trained observers, even though their observations may be of subjective character. Notice how we in education have sought measures which rule out the need for trained observation. We seek tests so objective that they can be scored by the janitor of the schoolhouse. We tend to be skeptical of any device which calls for the least bit of professional skill in its interpretation.

If in turn we examine the American educational program we will be impressed with the urgent need for the development of a measurement program. The fact that we are attempting to educate the masses puts us in need of the means whereby we can determine the educational characteristics, needs, and progress of every individual. If quality of living, rather than knowledge is the end point of education, we must have the methods for appraising quality of living both in school and ordinary life situations. This is especially true if we assume that the way to learn to live life on a high plane is to live life on a high plane. We have, to summarize, a need for three types of measurement, techniques for the study of individuals, techniques for the analysis and appraisal of learning activities and methods of measuring human conduct and living.

Techniques for the Study of Individuals

It is, of course, true that we are already in possession of many useful techniques for child study. In our use of these techniques we are handicapped by a tragic incompetence on the part of teachers and administrators. This incompetence is responsible for much of the abuse in connection with the application of measurement techniques. If we again use an illustration from the field of medicine we find that a patient's blood pressure, for example, is significant as a means of diagnosis, largely as it is viewed in relationship to a number of other factors. In medicine this emphasis upon a total picture of a case is considered basic. In the words of Sir William Osler "it is more important to know what kind of man has a certain disease than it is to know what kind of disease a man has." This concern over the individual as a whole is largely absent in our inter-

pretation of educational measures. We speak of children with high I Q's or low ones as the case may be, when we should know the I Q is significant only when considered together with and in relationship to a large number of other items of information about the individual. Frequently the facts are misinterpreted or given too great importance because of ignorance, bias, or over-simplification.

Perhaps the greatest need in child-study practice is in the direction of broader studies which are continued over a long period of time. Careful observation of the child at home and in school will become an important measurement technique. Such observation is not wholly objective but if it is made by trained experts it will have great value. Paper and pencil tests have value but their results will be considered in the light of other data secured through observation. Careful recording of a cumulative sort will be important.

Techniques for the Analysis and Appraisal of Learning Activities

The activity movement in education appears to be spreading. To be sure such tests as we have do not seem to show that children who have been taught in an activity school have learned more arithmetic or facts in history than children who have been enrolled in a traditional drill school. Proponents of the activity school are, however, far from dismayed over this situation. They insist that better living and not knowledge is the goal of education. If there is to be better adult living, there must be better child living. The school is an important part of child life. How, then, can we be sure that the child's life in the school is lived on a high plane? Here paper and pencil tests will not serve us. What we need is an analytical and critical description of class-room activities. Such descriptions will be checked against the ideals of living which are basic to the educational procedure in a given school. Let us assume, for example, that the school seeks to teach originality and initiative. If the supervisor in such a school wants to know whether the school is actually stimulating originality and initiative he can, from a careful description of the activities in a particular class, learn the number of opportunities which children have for the development of originality and initiative.

As teachers we often feel that medical men have more exact knowledge and more objective techniques than we have in education. One thing is clear and that is that the physician is much richer in carefully described and recorded experience. Medicine cannot utilize controlled experimentation in many areas. It can and does accumulate a wealth of case histories with detailed descriptions of both objective and subjective observations. We need a similar wealth of case studies with careful observation of activities and their results.

Measures of Conduct and Living

It is questionable whether so-called exact measurement has served us well in the direction of measurement of conduct and living. For one thing, it appears to have blinded us somewhat to the value of simpler procedures. In reading, for example, one may with complicated apparatus study

the eye-movements of a child. One may then laboriously attempt a remedial program. On the other hand, it is found that if children are given interesting reading material appropriately chosen with reference to their maturity, they will tend to develop the proper or most effective eye-movements and sometimes correct those which are faulty. Often we seek some statistical or mysterious measure with some almost occult virtue. In so doing we forget that ordinary common-sense observation scientifically interpreted may be all we need. The child who searches the shelves of the library every week for books on radio, does not need to be subjected to complicated interest tests in order to determine whether or not he is interested in radio. The boy who stays after school to study geography provides an illustration of the effectiveness of the work done by his teacher. So with countless other elements in human conduct. We recognize the various traits or elements of conduct in our friends and associates. We need but to be careful observers to note them in school children.

One reason for our failure as teachers to observe children more carefully is to be found in our general lack of interest in them. This lack of interest is especially striking on the high school and college level. Ask any high school boy about his opportunity for learning to know his teachers and you at once realize the need for a more human concept of education and the school. More can be learned about a pupil in a twenty-minute informal visit after school than in several months of classwork. It is severe reflection on us as teachers that we do so little to secure an intimate acquaintance with our pupils.

Summary

American education and its basic educational philosophy have developed along lines which challenge the underlying premises of our measurement procedures. In this way a serious conflict has arisen between our educational philosophy and our measurement principles. Unless this conflict can be resolved through the development of a program of measurement which is in harmony with our educational objectives, the whole measurement movement will be seriously retarded. This would be especially serious in America where, because of our program of universal education, we are in greatest need of child-study techniques. Here less emphasis upon objectivity and more upon trained observation would give us a more fruitful approach. Trained observation is also necessary in describing the quality of living in school and community. Paper and pencil tests alone are inadequate. Their results must be considered in relationship to a rich accumulation of data resulting from the teacher's life with the child over the entire period of his school and as far as possible his out-of-school life. The appropriate use of testing devices should not cause us to overlook ordinary common-sense observation. Such observation can be made far more significant through improvement in the teachers' professional knowledge and skill. The need for a high degree of professional intelligence is especially striking in the interpretation of all kinds of data descriptive of children and their learning activities.

LIMITATIONS OF FREEDOM

RIDGLEY C. CLARK
SUPERINTENDENT OF SCHOOLS
SEYMOUR, CONNECTICUT

THE most frequently used word in modern educational literature is "freedom". Its significance is often vague and unreal. It is surprising to note the variety of meanings that word may have and does have for different people. It may mean different things to any one person at different times. To some it connotes license to indulge whim or passion; to others it is an opportunity to work at an interesting task. There are restless souls who find it only in wandering to the remote places of the earth, contented souls who find it in their plain dirt garden, devout souls who find it in worship, valiant souls who in spite of prison walls or beds of pain find it in self-expression or the performance of some seemingly impossible task. In the full sense of the word, none is privileged to enjoy complete freedom.

Freedom is limited by physical and mental endowment. I attended a minor league baseball game some years ago and noticed that, as the batters took their places, their stance was similar to that of Babe Ruth. The form was that of Ruth, the performance was that of good ball players of less ability. They lacked by a fraction the quickness of eye, the perfection of timing, the power of swing that made this remarkable man the king of batters. Men have toiled long and zealously yet have failed to produce drama with the lasting appeal of Shakespeare, to write like Stevenson, to develop the delicate touch of Corot, to attain the mastery of finance achieved by the elder Morgan, to contribute to the field of invention like Edison. Their freedom to achieve is limited.

Many men are shut out by nature from certain fields of achievement. Call it heredity, call it what you will, they just are not the kind of person who succeeds in this particular field. The most famous preacher of his time is said to have been a failure as a teacher. Potentially good farmers have made poor preachers. Potentially good professional or business men have made only mediocre farmers. It is because that is true that we have introduced guidance into our school program. We are attempting to help pupils to select the field of their endeavor with some knowledge of what it is, and of their own fitness to the end that they may be more free to attain to the full measure of their capacity.

Freedom is limited by a lack of knowledge and skills. One of my favorite leisure-time activities is to go to a local shop and watch the skilled hammer men transform a rod of steel into the form of an auger or bit. To my amateurish observation the process is something like this: the steel is heated in the glowing furnace. When it has reached the proper temperature the skilled workman removes it, places it on the anvil, sets the automatic hammer in motion and by just the right turn of the arm and wrist so

manipulates the steel that, when he has finished the product has the form of a bit. The unskilled man cannot do this job. The operator has to know when the iron is at a proper temperature, how to start, stop, and regulate the speed of the hammer, how to adjust it and how to place the iron. To this knowledge must be added skill. To attain this skill there must be long hours of patient, intelligent practice. I suspect that, in the process of attaining it there is wasted material, discouragement, occasional blistered fingers. Possibly there is some profanity, but that is not necessary.

In trying to express ideas we are inhibited by lack of ideas, by the limitations of our vocabulary and our lack of practice in putting thoughts into words. We are not free to write as we wish. We need knowledge and skill.

To free the pupil to enjoy the delights of literature and art, to understand science, and to become a self-sustaining member of society, we open before him rich educational experiences, we teach him to read and give him an opportunity to know the classics and the best of contemporary literature and art. We teach him to read and work experiments in science. In a word, we try to make it possible for him to acquire knowledge. But knowledge is not enough. He must acquire skills. To do so involves drill. It is often asserted that progressive educators, so-called, do not believe in drill. This is not true. They, like other modern educators, remonstrate against dead and purposeless drill. Also, like other modern educators, they believe in purposeful, thoughtful drill in response to a felt need. Without drill the educative process is slow and ineffective. The more operations we perform automatically, the more time we have for activities that demand thought. My mother and yours too, if you are old enough, could rock a cradle, knit a stocking, and read a book all at the same time. This is an illustration of the freeing power of acquired skills. In offering study and practice, the public schools contribute to the pupil's freedom.

Lack of physical fitness holds back many from freedom. It is true that this is a real factor in the measure of freedom a person may enjoy. It is equally true that many fail to attain that measure of freedom which is their birthright. They fail to develop their bodily strength. The total amount of human energy that is wasted in combatting sickness is appalling. It impedes attainment.

A cultured gentleman, descended from negro slaves, stated that while the Emancipation Proclamation was a great boon to his race, a knowledge of sanitation which is checking the spread of tuberculosis has been equally instrumental in freeing the negro. The need of a knowledge of sanitation and of the rules of healthful living is common to all races. Nature's laws may be violated but not with impunity. Retribution may be slow or it may be swift, but it is inevitable.

Hence, the public schools attempt to instill into the minds of the young pertinent knowledge. Moreover, they attempt to inculcate habits that shall build up resistance to the invading enemies of bodily welfare, to produce stamina to meet life. Corrective measures are taken in the pre-school clinic. The school nurse and school

doctor cooperate. We have gymnasium and athletic programs. An oral hygienist in one of our great universities told me that each year the teeth of the freshmen seemed to be in better condition. She attributed it to the health program of the public schools. The purpose of the whole physical education program is to promote freedom.

Fears limit the freedom of many people. Even those far removed from childhood are afraid that the goblins will get them. Men in fairly prosperous circumstances live under a dread of poverty or poor health. Boys and girls are often shut out from the normal activities of childhood, such as climbing trees and playing leapfrog, because parents are afraid that they will get hurt. Many parents hamper the natural development of adolescent youth because they are afraid that in the process of learning to be social beings they will fall into moral delinquency.

Much fear is overcome by knowledge. In the presence of an eclipse of the sun the ignorant are afraid and hide in cellars; those who know smoke a piece of glass and gaze at one of the great phenomena of the universe. Yet life is dangerous. To save for a rainy day in times of prosperity, to satisfy our employer, to observe the rules of health, to teach our children to look both ways before crossing the street, to impress upon youth the fundamentals of decency is the part of wisdom. This the schools attempt to do. Such wholesome precautions are a part of wise living. Still life is dangerous and we read in school the stories of men who have lived it courageously even adventurously, who have met difficulties and reverses with determination. We try to develop an enthusiasm for the glorious adventure we call life.

Freedom is limited by the fact that we are social beings. We can keep a pig at our farm but not at our town house. Society may quarantine our family to preserve the health of others. Traffic rules must be obeyed. Property rights must be respected. The liberty of the individual must be sacrificed to the common good. Membership in society involves obligation and cooperation.

Society establishes good usage. The guest cannot wear his hat at dinner. Custom will not allow it. Our whole silly system of spelling is forced on us by custom. Some customs cannot be violated except under penalty of criticism or even ostracism by desirable people. Children, like their elders, hasten to conform. It is amazing how quickly a new style in dress, in haircut, in behavior spreads through a school. "Oh yeah?" became a part of every high school pupil's vocabulary almost overnight.

Schools teach and give opportunity for the practice of respect for the rights of others. They cultivate a concern for the common good. They teach cooperation and responsibility. This is the obligation that must be met before freedom can be enjoyed.

Yet there are vast possibilities for freedom within the social group. Important as wise conforming and cooperation are, the most useful members of society are not moulded by it. "Everybody does it" is an excuse, not a reason for doing wrong. One works as an artist, another as an artisan. One finds his recreation in books or

music, another in ten-mile hikes. Attainment is reached by those who strive while the mob lags. There is a glorious freedom for those who are not afraid of being laughed at because they believe and act differently. The group and society itself need leadership. Tradition is a great stabilizing force and as such holds society from disintegration. Blind allegiance to tradition prevents progress. New traditions must be developed if the race is to thrive. The school attempts to awaken the pupil to the possibilities of freedom, accompanied by responsibility. It tries to create a desire for individual rectitude and accomplishment which shall promote his own well-being and happiness and secure stability and progress for society.

Environment is defined as surrounding conditions, influence or forces. It is both geographical and social, a strong factor in shaping the lives of men. The highlander and the lowlander of fiction and fact have different characteristics. The loneliness of the rural districts gives opportunity for reflection while the city gives social grace. The peculiar forces that turn the rich toward snobbery, arrogance and profligacy or toward decency, responsibility, and philanthropy are far different from those that turn the poor to vice and shiftlessness or toward self-respect, and what are sometimes called the middle-class virtues. Some are hampered and helped by wealth, others by lack of it. The home in which the child is surrounded by attractive books and hears intelligent conversation has influence. The child who is in frequent contact with discord, crime or debauchery is bound to be influenced by it.

The school attempts to surround the child with an environment that is cheerful and wholesome. An atmosphere of friendliness pervades the activities throughout the day. An effort is made to break down some of the effects of unfavorable environment outside the school. The gifted in literature find that others gifted in music, art or arithmetic are equally honored and all take off their hats to the boy who finds it difficult to write a good letter but can buck the football line in a manner that excites admiration. It is learned at school that one need not be subjected to his environment entirely, but may rise above that which is unfavorable in it. Indeed, they may influence their environment.

The professional staff of the school is interested in outside agencies such as boy scouts, clubs, playground associations and churches that attempt to improve the environment of the child.

Men are bound by their past. What a person is today he is very likely to be tomorrow and a little more so. A boy or man sees a duty and puts off doing it. Later he sees other duties and delays. Finally, he forms a habit of procrastination and his efficiency is lowered. On the other hand, there are men who, when they see an opportunity to serve, seize it eagerly and perform the service. They become vigorous, capable men. He who has stolen is more likely to steal again. There are men who, on account of their past, find it difficult to do evil. They have built up habits of right living. Some have formed habits of success. So strong a force is habit

that it is much easier to form, than to reform. It is a function of the school to assist the pupil in building right habits.

Paradoxical as it may seem, to gain freedom one must sacrifice freedom. The student who attains honors may have to sacrifice freedom as far as other college activities are concerned. Under the NRA the manager is free to work longer hours than his men and to receive higher pay, but he must, however, assume responsibility. At the call of love young men sacrifice the freedom of bachelorhood tc become benedicts. The woman sacrifices the freedom of girlhood to satisfy her desire to bear and nurture children. If one would be free he must sacrifice freedom.

The schools try to teach these truths. They give opportunities for sacrifice in order to be free in books, and in social, vocational and physical fitness. They carry on that the young of this generation may find a large measure of freedom in their present and adult life.

SUMMARY OF PUBLISHED RESEARCH STUDIES IN THE FIELD OF LATIN TEACHING

LILLIAN DALE THOMAS
HEAD OF LATIN DEPARTMENT
FORT VALLEY, GEORGIA, HIGH SCHOOL

IN an article entitled "The New Spirit in Education," Claude M. Fuess, Headmaster of Phillips Academy, quotes Bishop William Lawrence as saying: "We studied our Latin and Greek grammar, learned interminable lists of exceptions to rules, and in spite, no doubt, of the best teaching of that time, those of us who were of only average ability got a distaste for the classical authors for whose enjoyment we had struggled through the grammar."[1] The Bishop and his schoolmates were by no means unique in this respect, but years of very fruitful investigation on the part of modern educators and psychologists are tending to make this "grammar poisoning"[1] experience more and more rare.

The present article does not attempt to raise again the moot question of the retention of Latin in the secondary school curriculum, but merely to review outstanding researches in the field of Latin teaching and to indicate briefly the major trends of these studies. According to historical sequence, the researches fall into three main groups: those made prior to the Investigation of the American Classical League, the Classical Investigation itself, and researches made since and, to some extent, based upon the Report of the Classical League.

In the group of early investigations, we find studies of Latin vocabulary, eye-movement and reading comprehension tests, investigations of the effect of Latin upon English and upon modern foreign language study, and surveys of a somewhat limited scope.

First of importance among the vocabulary studies was the word count

1 Fuess, Claude M., "The New Spirit in Education," Atlantic Monthly, 153:757, June, 1934.

of Dr. Gonzalez Lodge[2] based upon the selections from Caesar, Cicero, and Vergil usually read during the preparatory course. Dr. Lodge selected a basic vocabulary of 2000 words which he believed sufficient for the understanding of nine-tenths of the ordinary Latin which the student would be likely to read. The study was reprinted in the Teachers College (Columbia) Contributions to Education Series, and has furnished data for many more recent investigations.

Other early vocabulary studies were made by Elsie G. Hobson,[3] who compared two sets of Latin vocabulary tests to determine their efficiency in measuring both class and individual achievement and by L. W. Webb,[4] who found that the recall method of studying Latin words and phrases was superior to study with no conscious effort at recall during the study period.

The group of studies on ability to read Latin was initiated by the startling findings of the Judd-Buswell investigation of eye-movements in the reading of different languages. "The results show that in no case does a third-year student of the best grade in seven high schools in and around Chicago read Latin."[5] He appears rather to be engaged in some type of analytical study.

In 1924, B. L. Ullman and T. J. Kirby reported their effort to construct a reliable test of comprehension of Latin reading. They obtained a reliability coefficient of .71 $\pm$ .04 for eighth semester students.[6]

Several investigators centered their attention upon the problem of the effect of Latin upon English. In 1919, Wilbur Carr undertook to measure the pupil's ability to understand English words when he meets them. In six of the seven schools tested the Latin group was from 5 to 13 points better on the initial tests and in all schools from 12.5 to 26 points better on the final tests. It was also learned that "where special and definite instruction in derivation had been given [in Latin classes] the students did better on the tests."[7]

The effect of the study of Latin on ability to define English words was studied by Gilliland, who reported that when tests were given to 115 college freshmen "not only were those who had studied more Latin able to define more words of Latin origin, but they were also able to define more words of Anglo-Saxon and Greek origin appreciably better."[8]

Jacob Orleans made a study of the possible transfer value of the study of Latin to English vocabulary.[9] The judgments of competent persons were used as data and tables showing derivation of English words from Latin were compiled. The investigator real-

2 Lodge, Gonzalez, "Vocabulary of High School Latin" (Teachers College, Columbia University, New York, 1928).

3 Hobson, Elsie G., "Observations on Two Latin Vocabulary Tests," School Review, 28:509-517, September, 1920.

4 Webb, L. W., "A Comparison of Two Methods of Studying," School Review, 29: 58-67, January, 1921.

5 Judd, C. H., and Buswell, G. T., "Silent Reading": Chap. V, University of Chicago Supplementary Monograph 23, Chicago, 1922.

6 Ullman, B. L., and Kirby, T. J., "A Latin Comprehension Test," Journal of Educational Research, 10: 308-317, November, 1924.

7 Carr, Wilbur, "First Year Latin and Growth in English," School and Society, 14: 192-198, September 17, 1921.

8 Gilliland, A. R., "Effect of the Study of Latin on Ability to Define Words," Journal of Educational Psychology, 13: 501-505, November, 1922.

9 Orleans, Jacob S., "Possible Transfer Value of the Study of Latin to English Vocabulary," School and Society, 16: 559-560, November, 1922.

ized that the method was too subjective to yield very significant evidence.

Edward L. Thorndike reports in his study of the influence of first-year Latin upon range in English vocabulary that "Thorndike tests of word knowledge, forms A and B, were given in 1921-22 to all pupils, Latin and non-Latin alike, in the first year of 56 high schools," and that "the Latins made in general over one and one-half times the gain which the non-Latins of equal initial score made."[10] Thorndike also reported a study of the effect of first-year Latin upon the knowledge of English words of Latin derivation. "Pupils were given three forms of the Carr English vocabulary tests at the beginning, middle, and end of the term. The study shows a much larger gain for the Latin than for the non-Latin pupils."[11]

In addition to studies of the effect of Latin upon English vocabulary acquisition, Thorndike sought to determine the influence of Latin upon ability to read English. He found in his experiment that "the gain made by the Latins is about one and one-half times as great as the gain made by the non-Latins."[12] A rather sensational report of the effect of Latin translation upon English composition was made by Briggs and Miller in 1923. By analysis of over 300 translations from Cicero the authors conclude that Latin as generally taught harms English more than it helps. The study focuses its attention on the quality of English used rather than on the accuracy of translation.[13]

Two reports of the effect of Latin on modern language study are available. The first by Thomas Kirby analyzes the influence of previous Latin study upon the mastery of first-year French by 268 students at the University of Iowa (1921-22). Kirby found a low positive correspondence between number of years of Latin studied and marks in first-year French at college. He concludes that intelligence is a larger factor in earning grades than is previous study of Latin.[14] L. E. Cole investigated the relationship between amount of Latin previously studied and records of performance in first-year Spanish and college French at Oberlin College (1919-21). He also found a positive correlation.[15]

Two survey studies deserve brief mention. Carl Franzen reported in 1921 from Iowa State Department records involving 96% of all state high schools that Latin enrollments were five times those for any other foreign language.[16] W. L. Uhl obtained reports from 85,000 pupils distributed through all four years of the secondary school course and stated in 1924 that the average daily amount of time outside class devoted by Latin students to preparation of lessons was

10 Thorndike, Edward L., "The Influence of First Year Latin Upon Range in English Vocabulary," School and Society, 17: 82-84, January 20, 1923.
11 Thorndike, Edward L., "The Effect of First-year Latin Upon the Knowledge of English Words of Latin Derivation," School and Society, 18: 260-270, September 1, 1923.
12 Thorndike, Edward L., "The Influence of First-year Latin upon Ability to Read English," School and Society, 17: 165-168, February 10, 1923.
13 Briggs, Thomas H., and Miller, G. R., Jr., "The Effect of Latin Translations on English," School Review, 31: 756-762, December, 1923.
14 Kirby, Thomas J., "Latin as a Preparation for French," School and Society, 18: 563-570, November 10, 1923.
15 Cole, L. E., "Latin as a Preparation for French or Spanish," School and Society, 19: 619-622, May, 1924.
16 Franzen, Carl G. F., "Foreign Language Teaching in High Schools of Iowa," School Review, 29: 610-616, October, 1921.

considerably greater than was required for any other subject.[17]

From the entire group of investigations made before the survey of the Classical League the following general conclusions were drawn: (1) that a Latin vocabulary of a more or less limited scope should be mastered by the recall method of study; (2) that secondary school students do not read Latin, but rather engage in some process of analysis as shown by eye-movement photographs; (3) that comprehension of Latin read is difficult to test; (4) that especially where definite instruction to that end is given Latin students excel non-Latin students in their knowledge of English vocabulary; (5) that the study of Latin seems to increase ability to read English; (6) that the study does not improve the quality of a pupil's English composition; (7) that there is a positive correspondence between number of years of Latin previously studied and marks in first-year modern language work in college; (8) that enrollment in Latin classes exceeds that in any other foreign language group (Iowa high schools 1921-22); (9) that students spend a larger percentage of time in preparing Latin assignments than they spend on any other secondary school subject.

Many, if not all, of these conclusions needed substantiation by means of more experimental evidence, but they, at least, pointed the way for further and more thorough study.

In 1924, the Advisory Committee of the American Classical League published the *General Report* of their investigation based upon data from 750,000 tests given to 150,000 pupils in 1,313 schools, from questionnaires answered by 1,150 teachers, and from many special studies. The Committee recommended that "progressive development of the power to read and understand Latin" be considered the primary objective in Latin teaching. Many other objectives were held to be valid, such as: increased ability to read, speak, and write English; increased ability to learn other foreign languages; development of correct mental habits; development of historical and cultural background; and development of right social attitudes.

In regard to courses of study, the following changes were recommended: postponement of some and omission of other elements of grammatical content of the first year; reduction in amount and greater variety of kind of reading material; inclusion of early "made" or "adapted" Latin and postponement of reading from the first classical author until the fourth semester.

In regard to method, the following suggestions were made: comprehension of Latin should be taught in the Latin work order; much more attention should be given to understanding the thought content of the Latin read, and less to translation; translation, when used, should follow comprehension of the thought; new forms, syntax, and vocabularies should be met first in appropriate context; drill should be functional rather than formal; syntactical questions should precede rather than follow comprehension of Latin to be read; form drill should be directed toward the prompt recognition of the various possibilities of a given form.[18]

17 Uhl, W. L., "How Much Time for Latin?" Classical Journal, 20: 215-221, January, 1924.
18 American Classical League (Advisory Committee), "The Classical Investigation" (Princeton: Princeton University Press, 1924).

The Report of the Classical Investigation Committee was received with varying degrees of enthusiasm, censure, and indifference. It seemed, nevertheless, to stimulate much discussion and a wealth of further research in the field of Latin teaching.

The very next year (1925) Maxie Woodring reported an analysis of Latin and English examination books from the College Entrance Examination Board. The data presented in this study show that correct and effective English is not used in a large per cent of the Latin translations written by graduates of high schools who have had four years of Latin.[19]

An interesting experimental study of a visual method in Latin instruction was carried on by Fenton and Hill, who found that with their two very small groups of nine students each the new method was vastly superior.[20]

In 1928 Orleans and Solomon reported a study in Latin prognosis and the construction of a test to predict ability in the subject. The test battery is designed to measure certain specific abilities believed to be essential for success in Latin. The authors claim a correlation coefficient of about .80 when test scores are correlated with an average of teachers' marks and the New York Latin Achievement Test.[21]

Mark E. Hutchinson examined the results of a standard test in Latin syntax to see how difficulty and frequency of occurrence correlate. He found correlation to be low and recommended that lists of "Latin syntax demons" be made.[22] Another study of syntax was reported by Leo Kennedy (1932) who investigated the retention of certain Latin syntactical principles by first-year and second-year Latin students after various time intervals. He found that "identical [repeated] errors break down under the influence of study about three times more rapidly than identically correct responses are built up," and that "time is more effective in destroying identically correct responses than identically incorrect ones."[23]

A detailed analysis of the task of learning to read Latin was made by Clara Book "to determine what problems are involved . . . and to ascertain other facts about the learning process."[24] As this is an analytical and psychological discussion, the conclusions are not based on experimental proof.

Latin vocabulary has been the subject of many investigations. One by Anderson and Jordan is a study of the curves of forgetting for English equivalents of Latin words. There is found to be a high correlation between immediate memory and delayed recall, and a high correlation between intel-

19 Woodring, Maxie N., "A Study of Quality of English in Latin Translation," Teachers College Contributions to Education No. 187, Bureau of Publication, Teachers College, Columbia, New York, 1925.

20 Fenton, Norman, and Hill, Victor, "Experimental Study of Visual Method," School Review, 36: 675-678, November, 1928.

21 Orleans, J. S., and Solman, M., "Study in Latin Prognosis," Classical Journal, 24: 174-190, December, 1928.

22 Hutchinson, Mark E., "The Correlation between the Difficulty of Latin Constructions and their Frequency in High School Latin," Classical Journal, 24: 412-420, March, 1929.

23 Kennedy, Leo R., "The Retention of Certain Latin Syntactical Principles by First and Second Year Latin Students," Journal of Educational Psychology, 23: 132-146, February, 1932.

24 Book, Clara D., "Detailed Analysis of the Task of Learning to Read Latin," Journal of Educational Research, 25: 178-196, March, 1932.

ligence and each of the factors—immediate memory and delayed recall.[25]

The conclusions of the Carr study of vocabulary density in high school Latin present a potent argument for the preparation of reading material which is really easy as to vocabulary. The investigator found in the first 29 chapters of the *Gallic War* a vocabulary density of one to four and in Lhomond's *Viri Romae* a density of one to six. A pupil who has mastered only the present official word list for the first year will in reading the first 29 chapters of the *Gallic War* meet 453 familiar words and 518 words which he has never seen before. The case would be similar for a pupil reading *Catiline I* or *Aeneid I* with only the vocabulary preparation required by the graded word lists.[26]

Data "concerning the reliability of photographic eye-movement records in the reading of Latin and English, and concerning the validity of various measures of eye-movements in the reading, as shown by their intercorrelations and by their correlations with tests of comprehension . . ., with perception tests, and with general intelligence tests," have been assembled by Olivia Futch (1934) experimenting with 27 ninth-grade students. "There seems to be no significant relationship between any measure of eye-movements in the reading of Latin and comprehension of Latin. The data show that records of eye-movements may not be used alone as measures of Latin achievement. There is a very substantial relationship between eye-movements in the reading of English and comprehension of Latin. . . . A positive correlation (.52 ± .09) between IQ and number of regressions per line in reading Latin shows that the more intelligent students tend to make more regressions per line than the less intelligent. Eye-movement records furnish valuable information concerning the process of learning to read Latin, and aid in the diagnosis of difficulties. General intelligence is more closely correlated with comprehension of Latin than with eye-movements in the reading of Latin."[27]

The effect of the study of Latin upon character traits was the subject of an experiment by Elizabeth Meek, who used an experimental group of 20 tenth-grade students and a control group of the same number. Three tests were administered to both groups: Hill's Test in Civic Attitudes, Thurstone's attitude toward War Scale, and the Neumann, Kulp, Davidson Test of International Attitudes. All three of these measures consistently suggested "the possibility of developing desirable character traits by stressing them in connection with the teaching of Latin."[28]

Work with an advanced Latin section of 24 superior students was reported by J. F. Gummere of the William Penn Charter School, Philadelphia. At the end of the four-year period, the class was fully one year ahead of the regular sections and had done superior work all through the course.[29]

25 Anderson, J. P., and Jordan, A. M., "Learning and Retention of Latin Words and Phrases," Journal of Educational Psychology, 19: 485-496, October, 1928.

26 Carr, W. L., "Vocabulary Density in High School Latin," Classical Journal, 29: 323-334, February, 1934.

27 Futch, Olivia, "The Reliability and Validity of Photographic Eye-movement Records in the Reading of Latin," Journal of Educational Psychology, 25: 620-629, November, 1934.

28 Meek, Elizabeth B., "The Effect of the Study of Latin upon Character Traits," Journal of Educational Sociology, 7: 241-243, December, 1933.

29 Gummere, John F., "Four Years' Progress with an Advanced Latin Section," Education, 54: 617-621, June, 1934.

The bearing of Latin training upon the quality of other academic work in college was studied by C. H. Whelden, Jr., at Yale in 1930. He concluded that training in Latin "would help a man to do work of better quality as measured in terms of grades received in certain specialized courses . . . wherever the subject matter of such courses was directly tied to the knowledge gained in a study of Latin or where the background of knowledge gained in the study of Latin could serve rather directly as a source of information on the subject-matter of those fields."[30]

Several survey studies have been undertaken. Van Dyke compared the programs of study of 35 high schools at three different periods. In the period from 1906-11, 34 schools were offering Latin and in the language group it shared first rank with German, also offered in 34 schools. In the 1915-18 period, Latin had reached first place, with German second. In 1929-30, all 35 schools were offering Latin, which still held first place, with French second (offered in 33 schools), Spanish third (offered in 20), German fourth (offered in 18), and Greek fifth (offered in only 5).[31]

A survey was made by Land to determine the effect of the Classical Investigation upon Latin courses in schools preparing for college. Circular letters and questionnaires were sent to 125 schools which had prepared at least 100 candidates for the College Board examinations during the period 1926-1930. 52% of the schools sent complete returns. "It is only in the second year [course] that the effect has been noticeable. In this year 43% of the schools read less than the old requirements. In third and fourth years . . . only 17% and 13% respectively read less. . . . The schools which prepare the greatest number of candidates have been affected least by the Investigation."[32]

The National Survey of Secondary Education (1933) is familiar to all students of the subject. The monograph on "Instruction in Foreign Languages" was prepared by Helen M. Eddy, who, after examining many courses of study, visiting 82 Latin classes, and making inquiries by letter and personal conference, concludes: "The newer courses of Latin embody the recommendations of the Classical Report as to objectives, content, and to some extent materials, but the methods suggested by the Report and in course outlines which purport to lead to the attainment of the ability to read Latin as Latin are not being followed to any appreciable extent. Within the period since the Report the courses have been revised. The preparatory period has been prolonged, with much emphasis on early reading. The first classical author, Caesar, is postponed until the fourth semester, and the Cicero year contains selections from his letters and philosophical works as well as his orations."[33]

Researches in the field of Latin teaching have been many and varied.

30 Whelden, C. H., Jr., "Training in Latin and Quality of Other Academic Work," Journal of Educational Psychology, 24: 481-497, October, 1933.

31 Van Dyke, George E., "Trends in the Development of the High School Offering," School Review, 39: 737-747, December, 1931 (see also 39: 657-664).

32 Land, George A., "The Effect of the Classical Investigation upon Latin Courses in Schools Preparing for College," Classical Journal, 28: 179-186, December, 1932.

33 Eddy, Helen M., "Instruction in Foreign Languages," Office of Education Bulletin, 1932, No. 17, National Survey of Secondary Education Monograph 24 (Government Printing Office, Washington, D. C., 1933).

Much has been done, but much more remains to be done. In an article entitled "Some Needed Research in the Teaching of Latin," Mark E. Hutchinson briefly suggests several problems of major importance upon which investigations need to be made.[34] The whole question of prognosis and guidance, to which slight reference has been made in this discussion, demands further study, as do the problems of aim, content, and method. Too few of the investigations have been carried on by Latin teachers themselves. There is a place for such teachers, thoroughly conversant with research literature in the field, skilled in experimental technique, and capable of carrying on experimentation within their own classrooms. If Latin is to retain a place in the crowded high-school curriculum, it must do so on the ground of scientific evidence demonstrative of its value, and not because of bias on the part of its proponents.

34 Hutchinson, Mark E., "Some Needed Research in the Teaching of Latin," Classical Journal, 29: 335-356, February, 1934.

OLD PEDAGOGUES IN NEW PASTURES

L. F. Rover

On page 170 in this issue of Education, Dean Melby comments upon the "more or less general disappointment with the results of our measurement efforts." This article, written of course under an assumed name but by a professor in one of our outstanding universities, indicates at least one of the reasons why the measurement movement is criticized. Any help for Professor Rover sent to the office of Education will be forwarded to him.—H. B.

WHEN a man has been a pedagogue for twenty-five years he has seen changes occur in that vague thing we call a college education. He has been trying all his life to change people. Yet he himself is puzzled at some of the principal changes which he has observed about him. Do they mean progress? Or just the opposite?

I am one of these bewildered and confused pedagogues. My problem is an interesting one, and I state it here, hoping that some straight-thinking layman (or fellow pedagogue) may shed some light on it.

In the good old days a person who had a liberal education and who was really a superior scholar was sometimes fortunate enough to win the highest degree given in course, Doctor of Philosophy. Such a man naturally had some considerable acquaintance with philosophy, and did have a "liberal education". In these new days the candidate for the Ph. D. degree in most cases has no acquaintance with philosophy at all, and in many cases does no studying in the fields of history and literature. He is a specialist. He is "practical". His field of "research" is very narrow. He knows little and cares less about the words and deeds and ideas of the human beings who have preceded him on this planet.

What I, as an Old Pedagogue wonder about is simply this: Does the modern Ph. D. thesis, with its bad English, its "quantitative measure-

ments", and its almost jazzy modernity, represent a change for the better or the worse?

May I ask you, Gentle Reader, to judge for yourself from a few samples of such theses in the realm of the social sciences. To start with, take this doctoral dissertation on "The Automobile. A Sociological Study".

Ideas and inventions (says the preface to this thesis) are released into the social process without much thought upon consequences although they invariably disturb the current forms of social control and cause much personal and social disorganization. . . .

What are the responses which society and its members evince upon the entry of a new social force? Can these responses be reduced to types which can be utilized as a frame of reference for the study of other cases?

Let me confess right here that I like words. I like the expression "frame of references". I liked (when I first heard it) the term "pattern of activity". It occurred in a Ph.D. thesis, and was the scientific name for washing dishes. I always did especially like that observation of Oliver Wendell Holmes regarding his poem, "The Chambered Nautilus"—"a few sweet words,—words that have loved each other from the cradle of the language, but have never been wedded until now."

In this thesis is something new in "wedding words." One chapter is devoted to the "Theory of the Road"; another to the "Sociology of Locomotion and Social Life." Here I find this important information: "The road regularizes and channelizes the energy expended in locomotion." But the high point of this thesis is undoubtedly reached in Chapter 8—the chapter dealing with the "Psychology of the Automobile."

Omitting metaphysical considerations, the human beings, singly or collectively, view themselves as ultimate points of reference to all phenomena. This homocentric attitude is not so much a matter of conscious pride as it is a fact of psychological economy. It is an integration into some sort of intelligent system what would otherwise be a haphazard and disconnected chaos of events.

Our next Ph. D. thesis deals with that hard-headed, stubborn old English banker, David Ricardo. It seems that Ricardo had some pretty plain statements to make about the doctrine of scarcity. He was hard-hearted enough to claim that the scarcer any good thing became, the higher its price.

This thesis is on the subject, "Sociological Implications of Ricardo's Economics." The engaging young man that wrote this thesis piques our curiosity by such phrases as these: "objective rendering of social activities," "social activities in their wholeness." Here is a sample of the young author's attack on his problem:—

His (Ricardo's) detractors have done almost as much to center attention upon his work as have his most faithful disciples. This is doubtless due to the fact that his forcefulness of thought gave an authoritative air to his work that prevented even those who shuddered at his conclusions from questioning the accuracy of his conclusions if his foundations were granted. . . .

What ever else may be said to the credit or discredit of the sociologists, they have been laboring for more than a generation to convince all students of human experience that the best of their results are merely tentative, until they are correlated into an objective rendering of social activities in their wholeness.

This thesis on Ricardo leads us naturally enough to another one in the field of economics and sociology, "So-

cial Areas in Schuyler County." Here we find ourselves immediately immersed in the "new literature" of rural life — and the "rural community" as the modern nomenclature has it. What is a rural community? Here at long last is the answer:

> The aim of this study is to test the hypothesis of the rural community. . . .
>
> The people who habitually patronize or prefer a certain center regularly may be mapped or located geographically with respect to the place where the center is located and the boundary line, connecting the locations of the homes of those people most distant from the center who patronize it regularly in any service, determines the service area of the center for the service.

In browsing widely in this new field of literature — the doctoral dissertation—I have noted that not only are "words wedded" in a manner peculiar to this cult, but also that a new technique is employed in getting at the subject matter. The technique now in vogue is known by members of the cult as "quantitative measurements," but by the uninitiated it is called plain "statistics." So we see very earnest young men and women laboring like Sisyphus with pathetic zeal, trying to weigh the imponderables in human relationships, trying to measure the incommensurables of social life. Who was it that said, "You cannot measure grief with the chemistry of a tear! You cannot express a mother's rapture with an algebraic equation?" Our next thesis illustrates both the new literary language and the new technique of quantitative measurements. This study is called "Standard of Life in a Typical Section of Diversified Farming." It is not the farmers' standard of living that is here the object of analysis, but his standard of life. The matter and manner of this thesis justify more extended quotations than space permits.

> The term, Standard of life, as used in this study, treats of the subsistence and the satisfactions of the people with whom it is concerned. It may be regarded as a measure of life in terms of the sum total of values enjoyed by the family as evidenced through the use of time in the satisfaction of wants for things both material, as food, clothing, and shelter, and spiritual, as education, music, and art. . . . While the standard of living holds economic or pecuniary values as basic, the standard of life holds the realization of the highest human values attainable as the ultimate criterion of satisfaction.

The youth who writes this thesis explains his method of "measuring the standard of life." In a selected rural territory he made a house-to-house canvass, bearing in his hands a "questionnaire" of many pages. At each home he asked the housewife the questions on his printed schedule, and noted down in writing her answers. Here are typical questions, showing the thoroughness and intimacy of the research: "How many corset covers did you buy last year?" "How many night gowns?"

Space forbids the reproduction here of the author's chart for "computed cost of living." This chart includes these quantitative measurements and in the following order:— Food — puffed wheat or rice, dried beef, karo. Clothing — (wife, other women and girls) waists, aprons, petticoat, union suit, corsets, corset covers, night gowns, kimonos. Advancement — camera, education, postage stamps.

Looking still farther afield in the "new literature," and widening our quest for "wedded words" in novel combinations, we come upon a dis-

sertation entitled, "Relation of Cities and larger Villages to Changes in Rural Trade and Social Areas." Here we learn that a map is not really a map but a "diagramatic presentation of areas"; an outline is not an outline but a "conspectus." In an ingenious way the writer of this thesis has divided all areas into two classes, buying and social. The buying areas include "hardware, fine clothes, work clothes," and so on. The social areas include "Sunday school attendance, physician, telephones, rural mail routes, banking." The author's realism is almost breezy:—

> Physician areas are somewhat whimsical in as much as they depend largely upon a personal factor.
>
> From the density of the city movie and fine clothes spots over the country it would appear that the village areas are experiencing the keenest competition in the latter.
>
> They (gas stations and stores) are open 24 hours a day and seem to handle most anything the people want. They also do a big business Thursday afternoons when the other stores are closed and naturally, of course, they attract a good deal of night trade.

Let the literary critic turn his attention to this new type of literature which has so suddenly burst into full flower in North America. With so many universities conferring the Ph. D. degree each year, the prospects are good for a continuation and even an increase in this "Ph. D. English," if we may so term it. As these samples show, the young Ph. D.'s are creating a literature which has broken away from all the fetters of the old conventional standards. This is its first distinguishing mark. Its second characteristic is its intensely practical nature. Utility comes first. Only yesterday I was reading a young woman's thesis on a problem in Home Economics, namely, "Analysis of Noodles of Known Composition." And why not, I thought, analyze noodles of *unknown* composition? Among the young men's theses for the doctor's degree I have observed these titles:— "Analysis of the Cost of Growing Potatoes," "Farm Motor Trucks," "The Economic Study of Farm Buildings." As illustrating the dominating position of utility in these studies, I may cite the table of contents of this last thesis: "Cow Barn; Feed Alleys and Litter Alleys; Manager; Cow Stalls; Water Buckets." The energetic young man that assembled these data and organized these facts in due time received his reward—doctor of philosophy.

Here are the practical conclusions copied verbatim from a Middle Westerner's thesis:— "The large receipts of cheese during June and July do not effect (sic) cheese prices as much as the heavy receipts effect (sic) butter prices. . . . In general the movements in pig iron production precede those of cheese prices by about six months."

In my own soul my question is unanswered: In the past 25 years the Ph. D. thesis has changed all right, but has it changed for better or for worse? What is the answer?

DEPARTMENT OF BOOK REVIEWS

Our American Heritage. From Wilderness to Nation. By *Lillian S. Coyle*, Teacher of Social Studies in Roosevelt Junior High School, Philadephia, and *Walter P. Evans*, Supervising Principal of William H. Hunter School, Philadelphia. 303 pp. Price $1.60. McGraw Hill Book Company, New York.

There is still a considerable difference of opinion as to whether the social studies should be presented through the fusion arrangement or by well organized courses in history, geography, civics, etc. Considerable can be said in favor of both. *Our American Heritage—From Wilderness to Nation*, by Lillian S. Coyle and Walter P. Evans, in some respects is a compromise between the two schemes. The main presentation is through history, with frequent and sensible correlations with geography, civics, etc. In Unit I, "How America Became a Land of Opportunity," there is a section devoted to the topic, "Geography Makes History." Civics receive considerable attention in Unit III, "The Making of the American Citizen." Yet the main theme of the book is the history of the building of the nation.

The book is divided into three main units, namely: (1) How America Became a "Land of Opportunity"; (2) How the West Was Won; and (3) The Making of American Citizens. Under each unit are several topics. While it is relatively easy to keep a time sense in reading the book, the treatment of the material is topical rather than chronological.

The book reflects the Dewey philosophy of learning by doing. With each topic considered are a series of problems and activities. There has been an attempt to grade the problems according to differing abilities of the pupils. There are problems that all should be able to do—sort of minimum requirements,—and then there are those that only the brighter youngsters can do. Possibly individual skill and speed can thus be used to their utmost. Of course if homogeneous grouping is used this procedure will be of little use as a measure of individual difference but it will still possess motivating values.

Another interesting device is the series of check tests. These are of the new type. They should not be looked upon as achievement tests, but rather as to the ability to comprehend the reading of the presntation.

At the end of the treatment of each topical division of the unit, a splendid reading list is provided. These books are neither too advanced nor too simple for ready comprehension by the children.

The illustrations are excellent. There are several well chosen maps, although their inclusion at the rear of the book may nullify their usefulness somewhat.

While the book is planned for use in the junior high school, it could be used with advantage in grades below that level.
—FRANKLIN C. ROBERTS, Boston University.

The Bookman's Manual. A Guide to Literature. By *Bessie Graham*. Fourth Edition, revised and enlarged. 713 pp. 1935. R. R. Bowker Company, New York. Price $5.00.

The basis of this book, we are told by Miss Graham in her admirable preface, was a course of lessons in book salesmanship given in Philadelphia. She adds:

"Of the many things which must go into the training of a bookseller this volume ventures to touch but one field, the books themselves and their editions. . . . This volume is an introduction to book knowledge with some discussion of editions and their makers. . . . Bookselling is an ancient calling and an interesting business for which increased training is needed if our present day is to be served adequately in its rapidly growing need of books."

The 43 chapters in this comprehensive volume cover a very wide range of topics. Some of those that strike the layman as of special interest are: Books About Words, Early English Poets, Drama-History, Criticism, Collections, Continental

Drama of Today, Russian Literature, British Fiction (four chapters are devoted to this), American Fiction (three chapters), English and American Humorists, English History, American History, Art, and Music.

The comments on authors are brief but illuminating. Take for example Elizabeth Madox Roberts, 1886. After giving a list of her eight published books, the editors, and the price of each, Miss Graham says:

"Elizabeth Madox Roberts came into sudden fame with the publication of her first novel. It was a very long and diffuse but beautifully told story of a mountain girl, her sensitiveness to beauty, her nobility and her hard life. 'My Heart and My Flesh,' with a stronger structure, was less able to win the attention of the public. 'Jingling in the Wind' takes its title from Chaucer and is a modern Canterbury pilgrimage, the stories told by passengers in a stalled motor bus. 'The Great Meadow' is a story of pioneers in Virginia and Kentucky in the days of the Indians."

Not only will the bookseller find such a volume indispensable, but the teacher of English or American Literature will find it of constant help in locating authors, their books, and the publishers, as well as in giving a helpful estimate of the relative importance of a thousand different writers.—EVERETT L. GETCHELL, Boston University.

Progress in English. Book II. Doubleday, Doran & Co., New York. 523 + 12 pages. 1935. Price $1.40.

The authors, both of DeWitt Clinton High School in New York City, of which Mr. Clark is Principal, and Mr. Knox is Chairman of the English Department, are real teachers and know how to infuse the pages of the book with their teaching skill.

In this attractive and admirably arranged text, it is clearly evident that only the essentials for correct, clear, and forceful writing have been discussed, and consideration of principles beyond the pupils' powers of attainment have been avoided. The emphasis has been placed on those common errors of speech and writing that betray the uneducated person. For this reason, the sections devoted to grammar and common errors is unusually well done.

Only teachers of wide experience could have compiled so clear and practical a text. The illustrative examples are largely taken from themes written by high school pupils, and are admirably selected.

The arrangement of the material is logical, and the text well written. Some forty pages are devoted to the subject of Exposition, perhaps the easiest form of discourse for the beginner. Then in order are treated the paragraph, the sentence, description, narration, and argumentation. Business and social letter-writing receives clear and adequate treatment, as does the essay, précis writing, and oral expression. The next 70 pages deal with Grammar for Use,—one of the sanest treatments of this difficult subject that has come to our notice for a long time. The book closes with a chapter on words, and on punctuation.

The clear and concise definitions, the many well-chosen illustrative examples, and the pleasing appearance of this new text will make a strong appeal to all teachers of High School English who are looking for a new and interesting text.—EVERETT L. GETCHELL, Boston University.

NEW EDUCATIONAL BOOKS

We acknowledge receipt of the following publications of interest to the readers of EDUCATION. Mention here does not preclude review in this or subsequent numbers of EDUCATION.

A Syllabus in American History and Problems of American Democracy for Secondary Schools. By a Committee of the New England History Teachers' Association, 1935. $1.00. D. C. Heath and Company.

The Silver Book of Songs For the Grades. By *Clella Lester Perkins, Ann Trimingham, Mary Strawn Vernon, Louis Woodson Curtis* and *M. Claude Rosenberry.* 25 cents. Hall and McCreary Company.

Directed Activities, II. A Workbook to Guide Pupils in Their Study of the World Around Us. By *Samuel R. Powers, Elsie F. Neuner* and *Herbert Bascom Bruner.* 1935. 40 cents. Ginn and Company.

Study Guide to Everyday Food Problems. A Workbook for Home-Economics Classes. By *Jessie W. Harris.* 1935. 48 cents. Houghton, Mifflin Company.

Exploring in Physics. By *Reginald J. Stephenson,* the University of Chicago. With Drawings by *Chichi Lasley.* 1935. $1.50. University of Chicago Press.

The Psychology of Christian Personality. By *Ernest M. Ligon,* Ph.D., Department of Psychology, Union College. 1935. 393 pp., cloth, $3.00. The Macmillan Company.

Tested Sales Letters. By *Herbert H. Palmer,* Associate Professor of Business English and Advertising, College of Business Administration, Syracuse University. 1935. Cloth, 530 pp., $5.00. McGraw-Hill Book Company, Inc.

The following publications have all been received from The Bureau of Publications of Teachers College, Columbia University:

Parent Preferences of Young Children. By *Margaret Simpson,* Ph.D.

Intra-Family Relationships and Pupil Adjustment. By *Theodore R. Myers,* Ph.D.

The Size of the Local Unit for Administration and Supervision of Public Schools. By *Alonzo O. Briscoe,* Ph.D.

Some Parent-Child Relationships. By *Marian J. Fitz-Simons,* Ph. D.

Some Relationships Between Supply and Demand for Newly Trained Teachers. By *Mary Elliff,* Ph.D.

A Measurement of the Secondary School as a Part of the Pupil's Environment. By *J. Thomas Wade,* Ph.D.

A Study of Play Selection in Women's Colleges. By *Sister Mary Peter Doyle,* O.P., Ph.D.

The History of Journalism Education in the United States. By *DeForest O'Dell,* Ph.D.

Experimental Studies in the Psychology and Pedagogy of Spelling. By *Alice E. Watson,* Ph.D.

Measurement of the Personality Adjustments of High School Pupils. By *Percival M. Symonds* in collaboration with *Claude E. Jackson.*

The Nature of the Supervision of Student-Teaching in Universities Using Cooperative Public High Schools. By *Ralph F. Strebel,* Ph.D.

A Study Outline for School Staff Meetings. By The Staff of the Lincoln School.

Directing Study of High School Pupils. By *Maxie Nave Woodring* and *Cecile White Fleming.*

Millions of Years in a Winter. By *Edna Bridge Leining.*

Basic Student Activities. By *Joseph Roemer, Charles Forrest Allen* and *Dorothy Atwood Yarnell.* 1935. 367 pp., cloth, $2.20. Silver, Burdett & Company.

Personal and Community Health. By *Clair Elsmere Turner.* Fourth edition. 680 pp., cloth, $3.00. The C. V. Mosby Company.

A Textbook of Physiology. By *William D. Zoethout,* Ph.D. Fifth Edition. 1935. 694 pp., cloth, $4.00. The C. V. Mosby Company.